ESD-TR-68-148

THIRD QUARTERLY TECHNICAL REPORT-
LARGE APERTURE SEISMIC ARRAYS (LASA)

March 1968

DIRECTORATE OF PLANNING AND TECHNOLOGY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Sponsored by:

Advanced Research Projects Agency, Washington, D.C.
ARPA Order No. 800

(Prepared under Contract No. AF19(628)-67-C-0370 by General Atronics Corp., 1200 East Mermaid Lane, Philadelphia, Pennsylvania 19118.)

LEGAL NOTICE

When U. S. Government drawings, specifications or other data are used for any purpose other than a definitely related government procurement operation, the government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

OTHER NOTICES

Do not return this copy. Retain or destroy.

ESD-TR-68-148

THIRD QUARTERLY TECHNICAL REPORT-
LARGE APERTURE SEISMIC ARRAYS (LASA)

March 1968

DIRECTORATE OF PLANNING AND TECHNOLOGY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Spansared by:

Advanced Research Projects Agency, Washingtan, D.C.
ARPA Order No. 800

(Prepared under Contract Na. AF19(628)-67-C-0370 by General Atranics Carp., 1200 East Mermaid Lane, Philadelphia, Pennsylvania 19118.)

FOREWORD

This research was supported by the Advanced Research Projects Agency. The Electronic Systems Division technical project officer for Contract F19628-67-C-0370 is Major Cleve P. Malone (ESLE).

We wish to acknowledge the very considerable support and assistance that have been provided during the course of this program by the Advanced Research Projects Agency, the Earth Sciences Division of Teledyne, Inc., the Electronic Systems Division Seismic Array Program Office, the Institute for Defense Analyses, the Lincoln Laboratory, and the Vela Seismological Center.

This technical report has been reviewed and is approved.

William Lauterbach, Lt. Col., USAF
Chief, Seismic Arrays
Directorate of Planning and Technology
Electronic Systems Division

ABSTRACT

Four topics are discussed in this report. The first topic is the detection of underground nuclear explosions in the presence of large natural events. Previous studies of this problem have been extended and partially confirmed by applying calculations to additional data and by considering the effect of strengthening certain assumptions. The use of a LASA for estimating epicenters by beamforming and other methods is the second topic considered in this report. Recent studies of this topic have employed computer-aided simulations with controlled amounts of additive noise and travel-time anomalies to examine the relative performance of different algorithms for epicenter estimation. Simulations performed to date have indicated that a beamsplitting technique using a DIMUS array (i.e., only one bit per sample on each seismogram) leads to nearly the same level of performance available from the full analog array, and both beamsplitting techniques perform significantly better than techniques based on individual time picks on each channel, especially in the case of poor signal-to-noise ratios. The third topic considered in this report is an analysis of detection and false alarm probabilities for a DIMUS array. Taking advantage of a strong analogy between this problem and a standard problem in coding theory, bounds and asymptotic expressions for these probabilities are described, and the signal-to-noise ratio loss of the DIMUS array over the corresponding analog array is analyzed as a function of several parameters. Finally, the previously developed automatic pP test has been re-examined with a view toward evaluating its performance and simplifying its implementation.

TABLE OF CONTENTS

LIST OF FIGURES

SECTION I

INTRODUCTION

This is the Third Quarterly Technical Report on Contract F19628-67-C-0370. The report summarizes progress on four topics: the detection of underground nuclear explosions in the presence of large natural events, beamforming and other techniques for epicenter locations, detection and false alarm probabilities for DIMUS arrays, and a re-examination of an automatic pP test.

Section II discusses the problem of detecting underground nuclear explosions in the presence of large natural events. The work described in this section is an extension of previously reported work. In our Second Quarterly Report, data were presented that indicated the sensitivity of continental-size arrays in detecting covert underground nuclear tests. These calculations were based on simulations using actual seismic data from a single large earthquake. The calculations have now been repeated using data from another earthquake. These results are presented in Section II and indicate that the previously reported statistics may be regarded as representative. In our First Quarterly Report, we suggested that a network of single seismometers would be useful in monitoring for a covert underground nuclear test. In that report it was assumed that if any of the single seismometers received the signal from an underground test prior to the signal from the earthquake, the test would be detected. Based on this assumption, minimum time delays were calculated as a function of the position of the test site relative to the earthquake epicenter. In this section, we consider the effect of strengthening this assumption to require that at least two or at least three of the single seismometers receive the test signal before the earthquake signal. The consequences of strengthening this assumption are examined for two collections of stations. Based on these calculations, it would appear that some upgrading of the quality of existing stations

would be necessary to provide a satisfactory network of single seismometers for this purpose.

Section III reports the results of continued studies on beamforming and other methods of epicenter location with a LASA. Our Second Quarterly Report described some theoretical analyses and simulation studies of this problem. The simulation studies had been continued, with some modifications, during this quarter and some additional theoretical analyses have been performed. It has been found that the automatic detector previously employed triggers earlier than desired on very large events, and, as a consequence, the location errors presented in the last report were occasionally larger than expected. The automatic detector has been modified, and numerous simulations have been performed with controlled amounts of additive noise and travel-time anomalies. Four different algorithms for estimating epicenters were used in these simulations. Of these, the beam-splitting techniques performed significantly better than techniques based on individual time picks on each seismogram whenever the signal-to-noise ratio was small. Of particular interest in this study is the performance level that has been achieved by the DIMUS beamsplitting algorithm, which uses only one bit per sample of the individual seismograms. In the case of low signal-to-noise ratios, this algorithm has yielded a level of performance almost equal to that of the analog beamsplitting techniques, and significantly better than that available by plane-wave fitting after individual arrival time measurements. When the signal-to-noise ratio is very high, of course, there is relatively little difference between the different algorithms. These results must be regarded as preliminary, since they are based on only one collection of seismic waveforms plus varying amounts of additive noise and time anomalies. However, it is expected that similar results will obtain when additional data are considered.

In Section IV, the detection and false alarm probabilities for a DIMUS array are analyzed for a variety of parameter values. A very strong analogy between this problem and a familiar problem

in coding theory is employed in order to obtain fairly tight bounds on detection and false alarm probabilities as well as asymptotic expressions for their behavior. These bounds and the asymptotic expressions provide relatively easy means of obtaining good estimates of these probabilities without having to evaluate cumbersome sums of binomial coefficients. These techniques are illustrated in several ways for various parameter values, and the performance levels available from a DIMUS array are contrasted with those available from the corresponding analog array. It is observed that in a limited range of parameter values the signal-to-noise ratio loss of the DIMUS array compared to the analog array is 2 dB; in general, it is somewhat larger. These results are discussed and illustrated with specific examples chosen to be relevant to present and future seismic arrays.

Section V presents a summary of a re-examination of the automatic pP test, which has been described in earlier reports. This study was undertaken to understand more clearly under what circumstances this test yields results that would otherwise be unavailable, to understand what irreducible limitations on the test performance do exist, and to consider what simplifications in the implementation of the test might be possible without a serious degradation in its performance. In order to consider these topics, several specific events have been reprocessed and the results are presented in this section. In addition, previous calculations have been reviewed and certain statistics abstracted from these. Based on these studies it is concluded that the use of very large arrays for enhancing the pP phase does yield results that are sometimes not available by individual examination of the seismograms. The adjustment for P-pP delay times appears to be a useful procedure in combining seismograms from these very large arrays. In the case of moderate depth earthquakes (10-40 km) it now appears that the fine-grained calculation of the average paired correlation coefficient may be unnecessary and that the same information

could be obtained by calculating this statistic on a 10-second interval, thereby reducing the number of computations necessary to carry out the test. Whether or not this conclusion applies when the test is used for depths greater than 40 km is not yet clear. Finally, it is observed that apparent reflections resulting from velocity discontinuities in the earth's crust may produce additional signals that cannot easily be distinguished from the legitimate pP and sP phases. It may well be that the only way to avoid ambiguities as a result of these reflections is to acquire adequate knowledge of the geology local to the epicenter region.

SECTION II

DETECTION OF UNDERGROUND NUCLEAR EXPLOSIONS IN THE PRESENCE OF LARGE NATURAL EVENTS

The problem of detecting an underground nuclear explosion that occurs shortly after an earthquake and in the same geographical vicinity has been considered in earlier reports [1,2]. A detection system consisting of a continental-size array plus a set of single stations encircling the earthquake epicenter at close ranges has been considered, and the properties of this hypothetical system have been studied. The single stations impose a minimum delay time between earthquake and test,and the continental-size array imposes a maximum. If either constraint is violated, the test risks detection. If the minimum delay time is violated, the signal from the test will reach at least one of the single seismometers before that of the earthquake. If the maximum time delay is violated, the earthquake signals will have decayed sufficiently that the continental-size array will be able to detect the test. Results of additional studies of both components of this detection system are reported in this section. Since this work is a continuation of earlier studies that are reported in our First and Second Quarterly Technical Reports, several references to these reports will appear in the following discussion.

In the Second Quarterly Technical Report [2], the dependence of the detection sensitivity of the array on the time after the earthquake and the location of the covert test relative to the earthquake epicenter was studied using seismic data from a single large earthquake recorded at fifteen stations located in the continental United States. In order to be sure that these results are representative, similar calculations have been repeated using data from another large earthquake and the results of these calculations are summarized below. Based on these additional calculations, it is concluded that the previously reported results may be regarded as representative.

The second component of the proposed detection system is a network of single stations encircling the earthquake epicenter at close ranges (perhaps 40°). Assuming that a covert underground shot would be detected if the signal from the explosion reached any of the single seismometers before the earthquake signal, a minimum delay time between earthquake and shot may be calculated for possible test sites in the vicinity of the earthquake. Results of these calculations were summarized in the First Quarterly Technical Report [1] in the form of minimum earthquake-test delay time contours on a map of the earthquake epicenter region. Recently, assumptions more realistic than the original one that a single station receiving the explosion signal before the earthquake signal would be sufficient to indicate the possibility of a covert underground test have been considered. In particular, the requirement that two, or even three, stations receive the signal from the explosion before the signal from the earthquake has been considered. With the small number of stations originally considered, this stronger requirement significantly enlarges the geographical area covered by a given minimum delay contour. If, however, a larger number of stations is assumed these contours can be reduced to a more reasonable size. This point will be illustrated in the discussion below by presenting calculations based on existing sites of present world-wide standard seimograph stations.

2.1 SENSITIVITY OF THE CONTINENTAL-SIZE ARRAY

Two detection statistics that are based on the records available from a continental-size array were proposed and their properties were studied in the Second Quarterly. These studies were based primarily on simulations employing seismic data obtained at fifteen stations in the continental United States from a magnitude-5.5 earthquake that occurred in the Fox Islands on 1 September 1964. The same calculations have been performed on seismic data obtained from fourteen stations in the continental United States from a magnitude-6.0 earthquake that occurred in

the Aleutians on 17 March 1965. Portions of the results of these calculations will be summarized here with references to the relevant figures of the Second Quarterly.

The two detection statistics under consideration may be defined as follows. The energy ratio detection statistic, ε, is defined as the average power in the array output in a one-second window divided by that of the preceding five seconds. The second detection statistic, η, is obtained by multiplying ε by the average-paired-correlation-coefficient obtained for the same one-second interval, using the individual seismic records. For both statistics the array output is obtained by aligning the seismograms to "steer" for a given "grid point," and calculating a weighted sum of the aligned seismograms. It has been shown [2] that the maximal-ratio weighting of the individual seismograms is superior to an equal gain weighting in this context, and, therefore, only this weighting is considered in this report.

In the Second Quarterly, the maximum excursions of the η statistics were tabulated for each grid point in the vicinity of the earthquake epicenter [see Figure 2.6 of Reference 2]. Based on these results it was inferred that a detection threshold of 2.0 would be a conservative threshold in terms of false alarms. The results of the same calculations performed on the magnitude-6.0 earthquake are presented in Figure 2.1, and it may be observed that the same conclusion still applies. As would be expected, there are several differences in the numbers resulting from the two earthquakes; however, the general range of values seems comparable, and does not appear to be dependent upon the magnitude of the earthquake.

In the Second Quarterly it was demonstrated that the η statistic could be expected to exceed the threshold of 2.0 whenever there was an energy excursion, as measured by the ε statistic, in excess of 10 dB. As a result of this observation, by simply examining the time dependence of the ε statistic for each grid point, it is possible to conclude how large an underground test would be necessary, as a function of time and test-

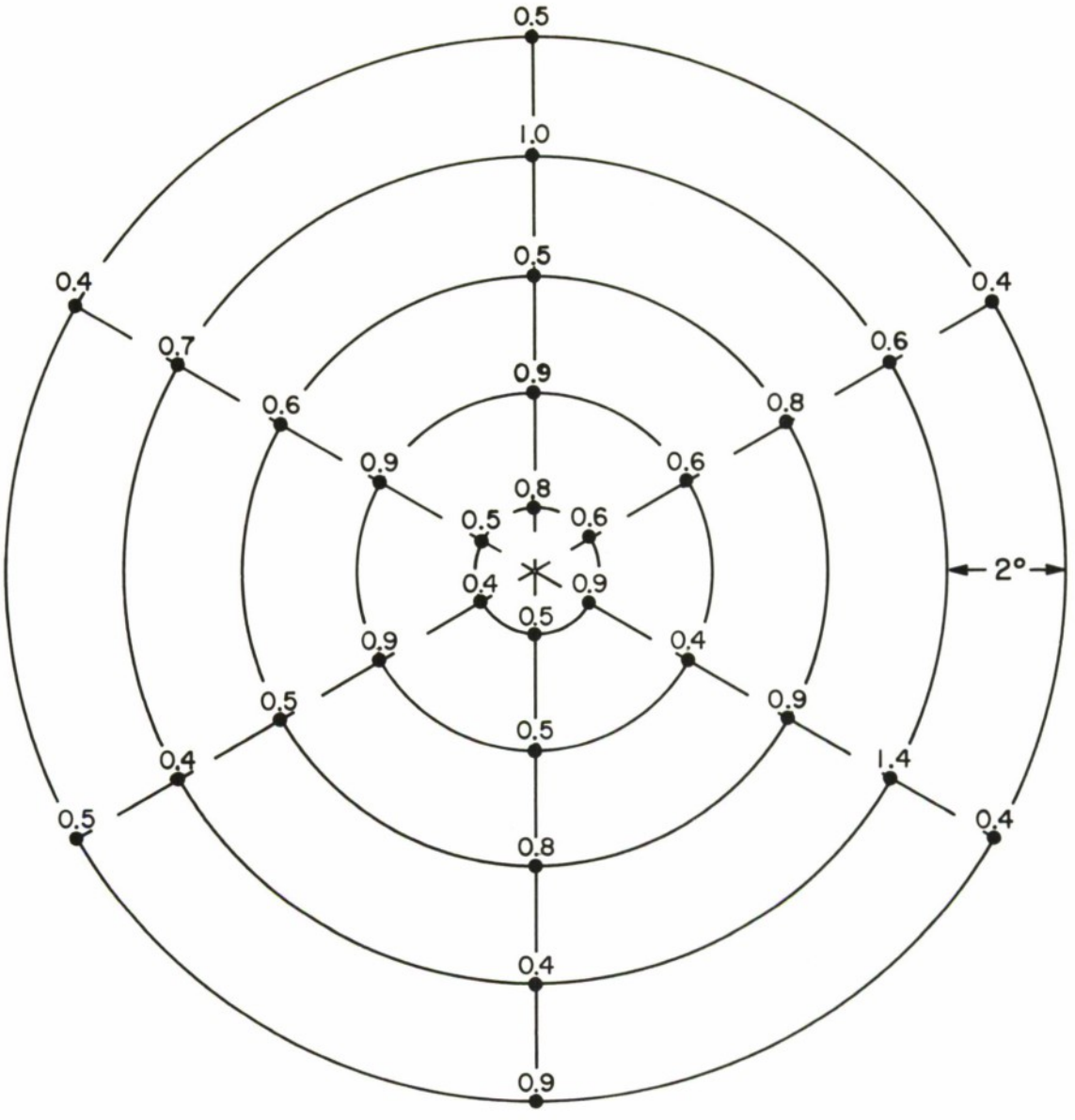
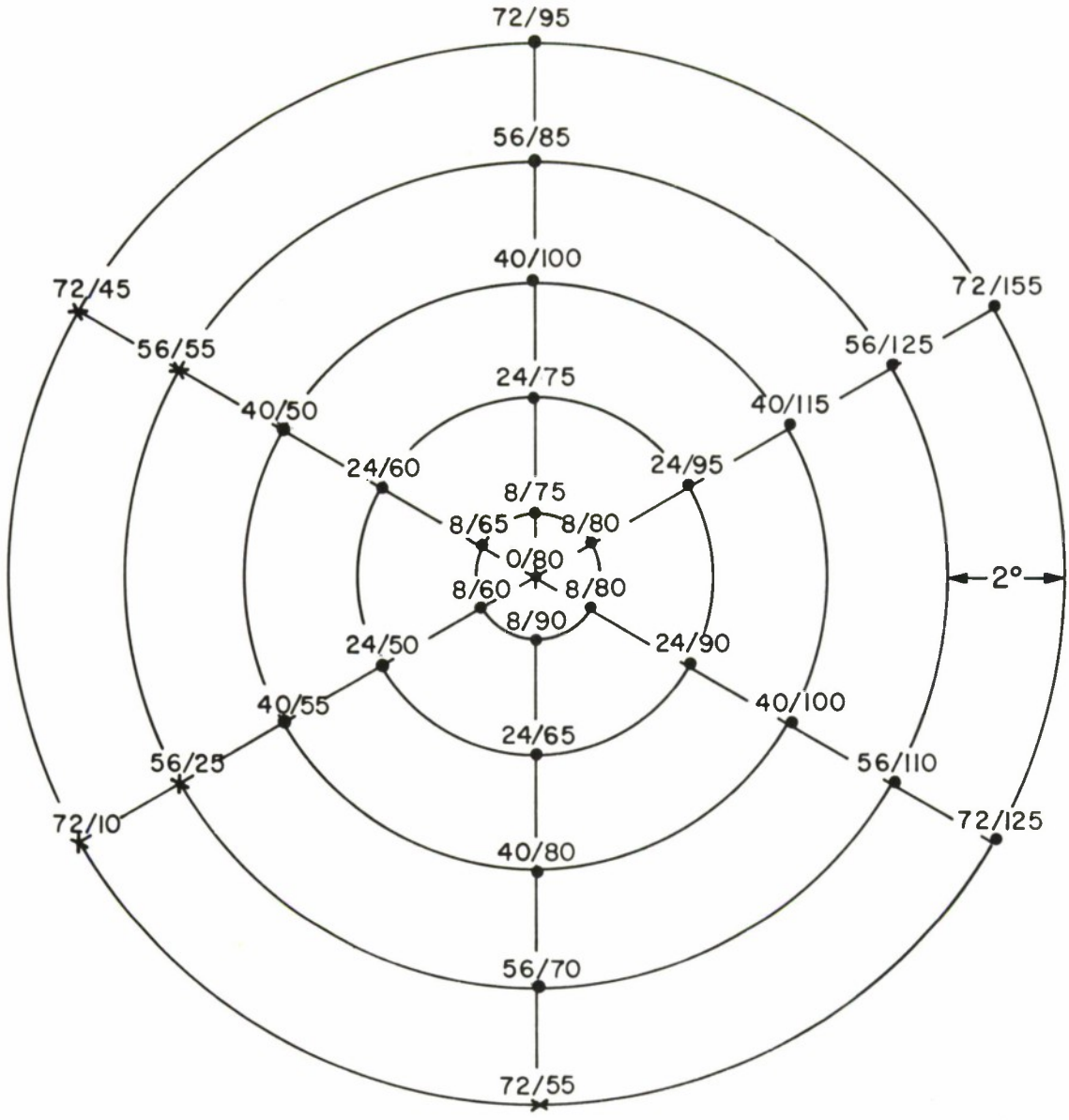


FIGURE 2.1
MAXIMUM EXCURSIONS OF η STATISTIC
EXCLUDING P-PHASE ARRIVALS

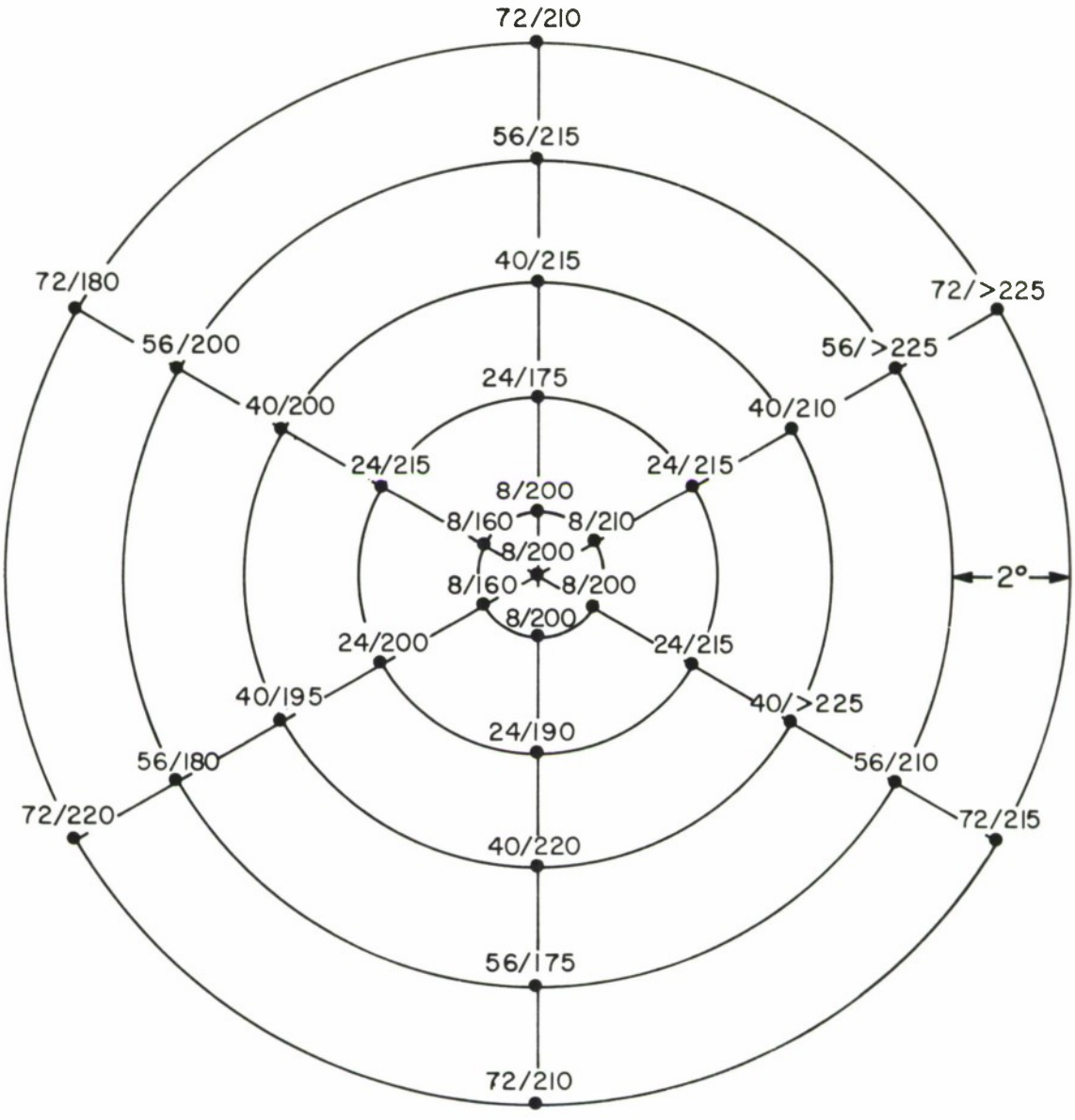
site position, in order to allow detection by the continental-size array. These calculations have been performed with data obtained from the magnitude-6.0 earthquake and the results are comparable to those which would be predicted by adding 0.5 magnitude units to the results of the calculations based on the magnitude-5.5 earthquake.

One way to summarize the detection sensitivity of the continental-size array is in terms of the maximum delay time between earthquake and underground explosion as a function of test site that is imposed by the continental-size array. Using data from the magnitude-6.0 earthquake, the time intervals for concealing a magnitude-5.5 and a magnitude-5.0 shot are presented in Figures 2.2 and 2.3. (These figures are comparable to those based on the magnitude-5.5 earthquake reported in Figures 2.12 and 2.13 of the Second Quarterly.) As before, two numbers are given at each grid point. The first is the minimum delay time imposed by a hypothetical network of single seismometers located at a range of about 40° from the earthquake epicenter. The second number is the maximum delay time imposed by the continental-size array. In the cases where the minimum delay time exceeds the maximum delay time, an X is indicated at the grid point on the figure to emphasize that an underground nuclear test of the assumed magnitude could not be detonated at that grid point without risking detection by either the single station network or the continental-size array. Comparing the time intervals available for concealment with those obtained earlier from different data indicates values that are roughly comparable, but by no means identical. Considering, for example, a nuclear test that is one-half magnitude unit smaller than the earthquake leads to maximum delay times based on the magnitude-6.0 earthquake data that are in most cases within a factor of two of those obtained using the magnitude-5.5 earthquake data. (Note that in both cases a shot that is one-half magnitude unit smaller than the earthquake is assumed, i.e., the shot magnitude is 5.5 in the first case and 5.0 in the second.) Similar observations



RECORDS NORMALIZED USING
MAXIMAL-RATIO COMBINING

FIGURE 2.2
TIME INTERVALS AVAILABLE FOR CONCEALING
(TEST MAGNITUDE 5.5)



RECORDS NORMALIZED USING
MAXIMAL-RATIO COMBINING

FIGURE 2.3
TIME INTERVALS AVAILABLE FOR CONCEALING
(TEST MAGNITUDE 5.0)

apply to the case where the shot is one full magnitude unit smaller than the earthquake. Based on these calculations, it is concluded that the results presented earlier may be regarded as representative and the continental-size array together with the network of single seismometers seriously limit the time interval during which a nuclear test could be concealed. As before, it should be noted that the number of computations implied by this detection network is very large, so that there still remains a serious practical question as to whether such a network could reasonably be operated.

2.2 SINGLE STATION DETECTION

In the First Quarterly it was assumed that if any member of a network of single stations received the signal from an underground test before that of the nearby earthquake, the test would not escape detection. Based on this assumption, minimum time delays between earthquake and underground nuclear test were calculated as a function of test site position relative to the earthquake epicenter. The calculations originally presented used a set of eighteen WWSSS stations and assumed an earthquake epicenter in Mongolia. The resulting minimum time delay contours are very roughly circular with a radius of 2°, 3.5° and 8° for minimum time delays of 15, 30 and 60 seconds, respectively [see Figure 3.2 of Reference 1]. If this original assumption were modified to require that at least two of the stations receive the signal from the shot before that of the earthquake, then these contours would grow appreciably. Assuming the same set of stations and the same epicenter, calculations show that the 15-second contour would stretch as far as 5 degrees in one direction and the 30-second contour as far as 10 degrees. Therefore, with this particular collection of stations, to require more than one station significantly increases the size of the minimum time delay contours.

If the network of single stations contained a great many stations that were well distributed in azimuth, an increase

in the requirement, from at least one station to at least two or even three stations detecting, would not seriously affect the contours. This observation is illustrated by the minimum delay-time contours given in Figure 2.4. These contours are based on a very large number (116) of stations from the existing WWSSS network. The 15, 30 and 60 second contours based on the original assumption that a single station detects are given in solid lines,and the contours implied by the requirement that at least three stations detect are given by dashed lines in Figure 2.4. As would be expected in this case of a great many stations, the region covered by these contours is moderate in size and not strongly affected by the more stringent requirement of three stations detecting. The stations used in this calculation and the epicenter under consideration are indicated in the map of Figure 2.5.

The collection of stations considered in the above calculations are unrealistic in that many may not be capable of detecting moderate size shots even without the earthquake interference. In a sense the results of assuming this complete set of stations represent an upper bound on performance which could be achieved with existing station sites. In order to obtain a more realistic measure of what level of performance could be obtained from existing WWSSS stations, the following exercise was performed. From the C&GS Earthquake Data Reports seven earthquakes with magnitudes between 4.7 and 5.0 that occurred in the Kuriles-Kamchatka region during 1966 were selected. Then 23 stations, each of which reported at least four of the seven earthquakes under consideration, were chosen. These 23 stations are indicated in the map of Figure 2.6. Using these 23 stations plus the same epicenter as before (in the Kamchatka-Kuriles region), contours of minimum delay time between earthquake and test were calculated for each of three assumptions: at least one station receives the explosion signal before the earthquake signal, at least two, at least three. The results of these calculations are presented in

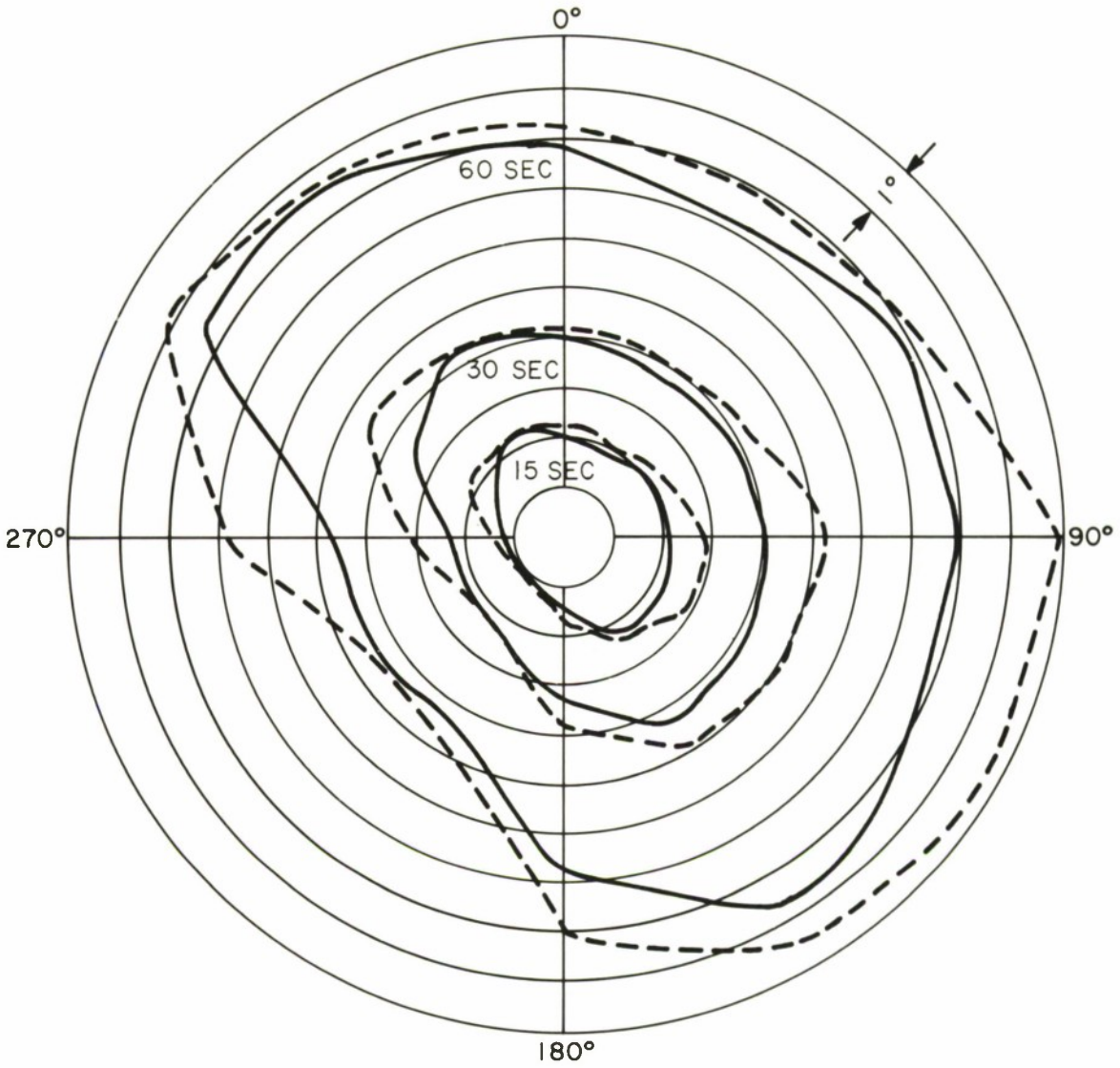


FIGURE 2.4
MAXIMUM TIME DELAY CONTOURS
116 STATION NETWORK

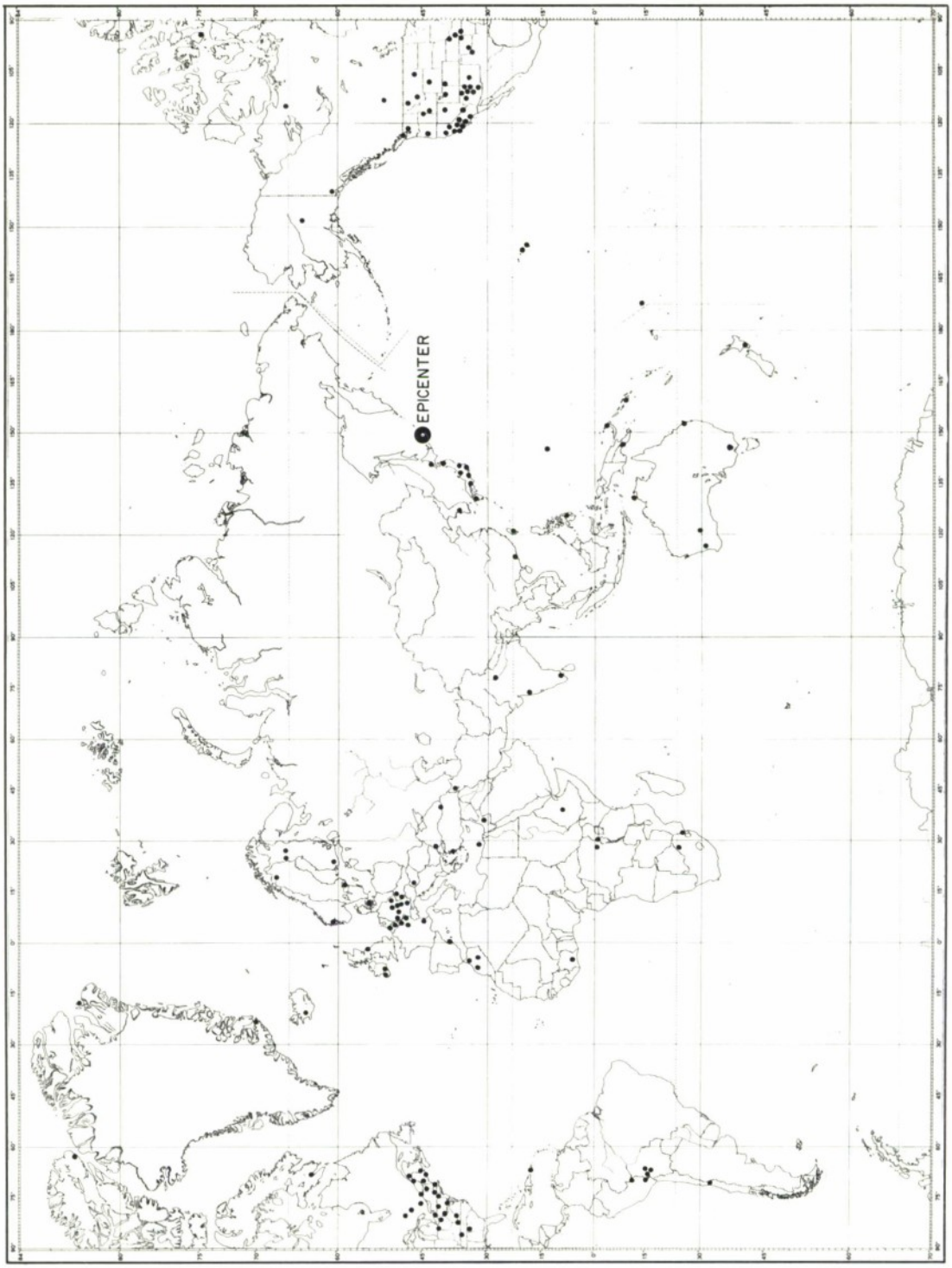


FIGURE 2.5 MAP OF 116 STATION NETWORK

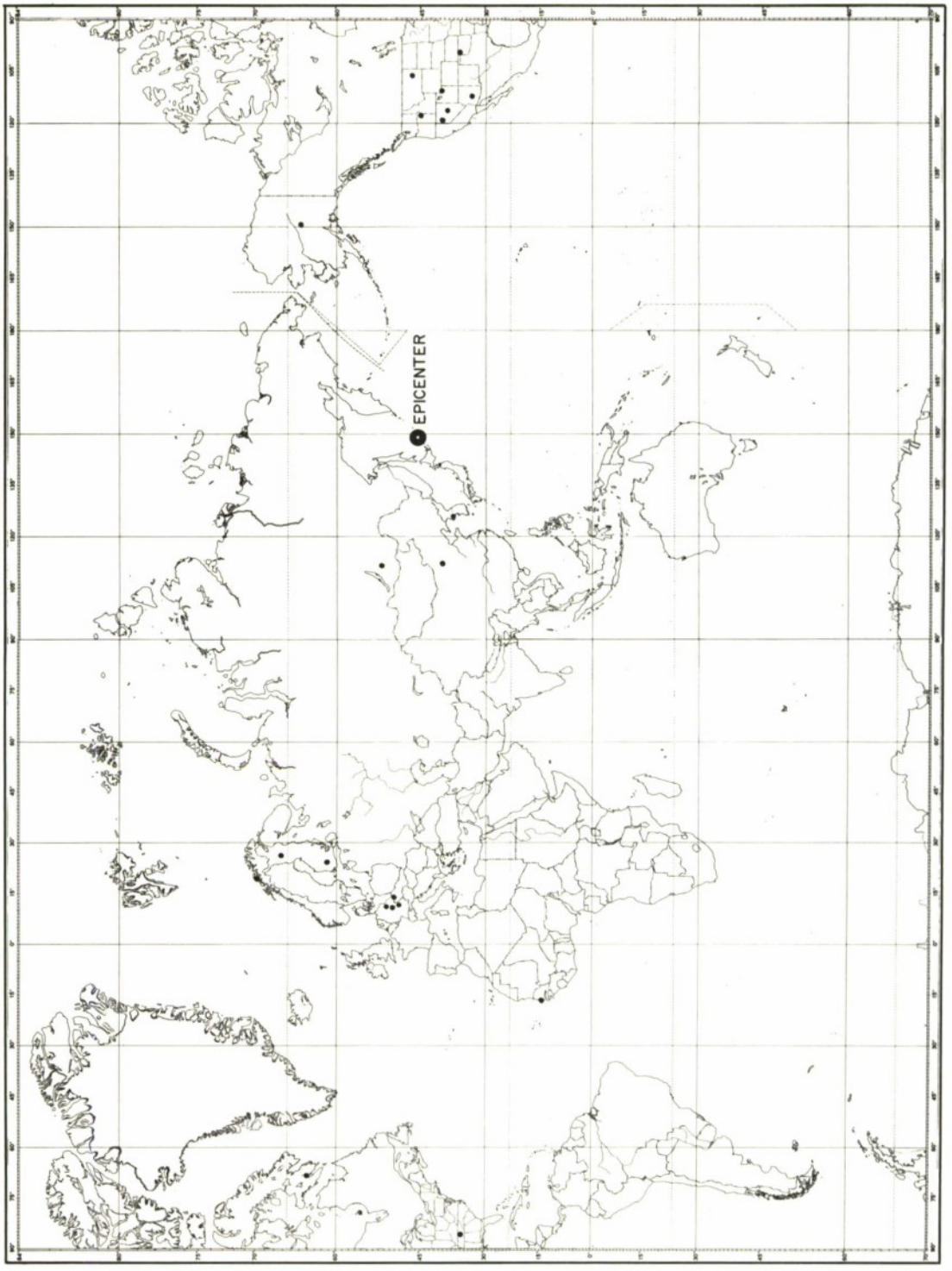


FIGURE 2.6 MAP OF 23 STATION NETWORK

Figure 2.7A-C. From this figure it may be observed, first of all, that these stations are not well distributed in azimuth and that therefore test sites bearing 150 degrees from the earthquake epicenter are not seriously affected by this single station network. Secondly, it may be observed in going from one to two stations, and especially in going on to three stations, the contours increase significantly in size. For both reasons it does not appear that this particular collection of 23 stations would be satisfactory as a single station detection network. In a sense, this collection of stations represents a lower bound on the performance that could be expected immediately with existing WWSSS stations. From this and the previous figure it would appear that a satisfactory network of single stations could be obtained using locations near those of the WWSSS stations provided adequate (single-station) detection sensitivity could be achieved at these locations.

2.3 SUMMARY AND CONCLUSIONS

Additional calculations relating to the problem of detecting covert underground nuclear tests have been presented. These calculations represent a continuation of earlier study on this problem in which a detection network consisting of a continental-size array plus a network of single stations well distributed in azimuth about the epicenter was proposed. In earlier reports the sensitivity of the continental-size array was evaluated by using actual seismic records from a single magnitude-5.5 earthquake. Additional calculations based on data from a magnitude-6.0 earthquake have been presented above and suggest that the original results may be regarded as representative.

Previously reported properties of the single station network have been reconsidered in the context of somewhat more realistic assumptions. The original assumption was that if the signal from the underground test arrived at any of the single seismometers before the earthquake signal, the underground

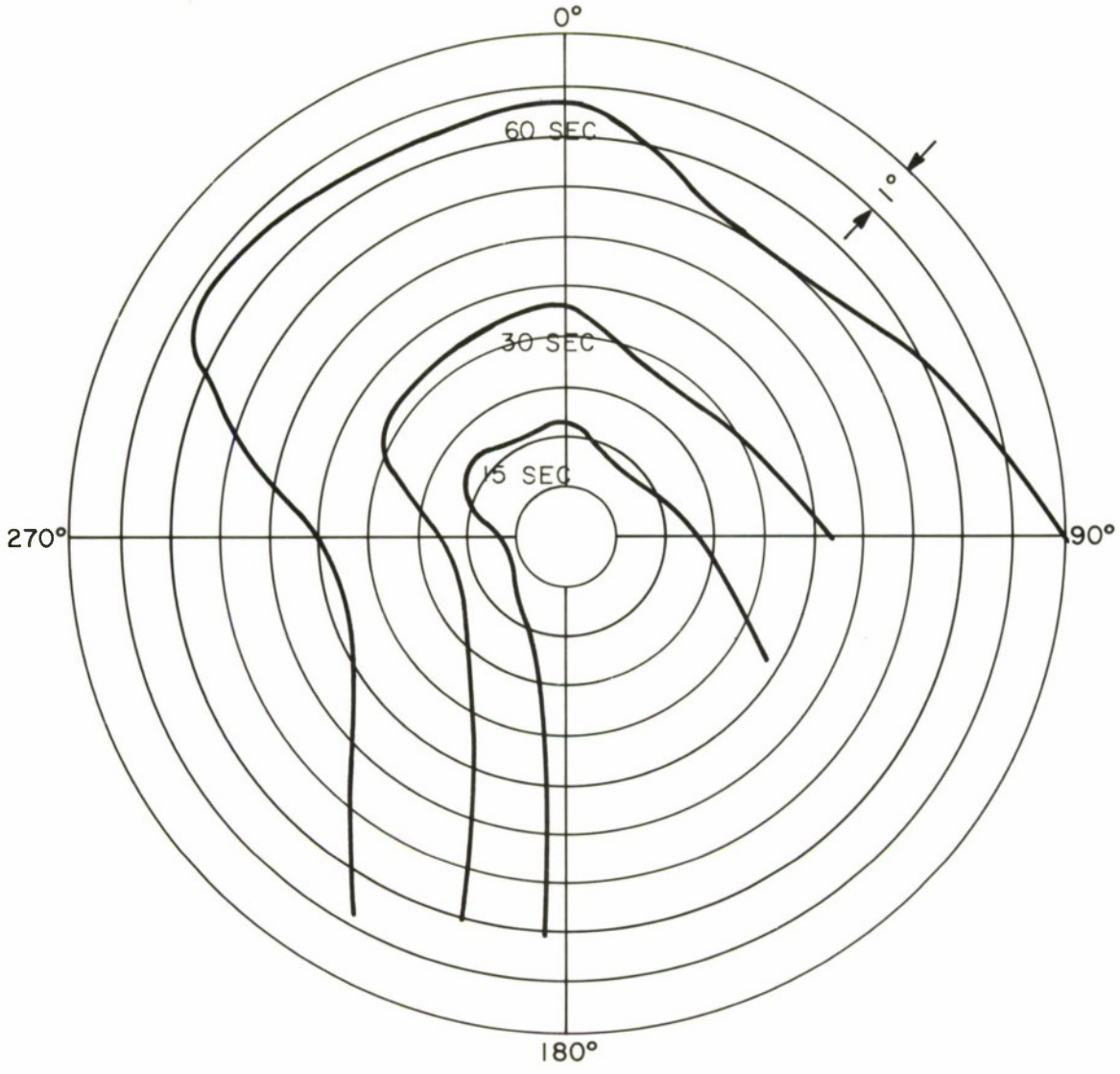


FIGURE 2.7 A
23 STATION TIME DELAY CONTOURS
SINGLE STATION CRITERION

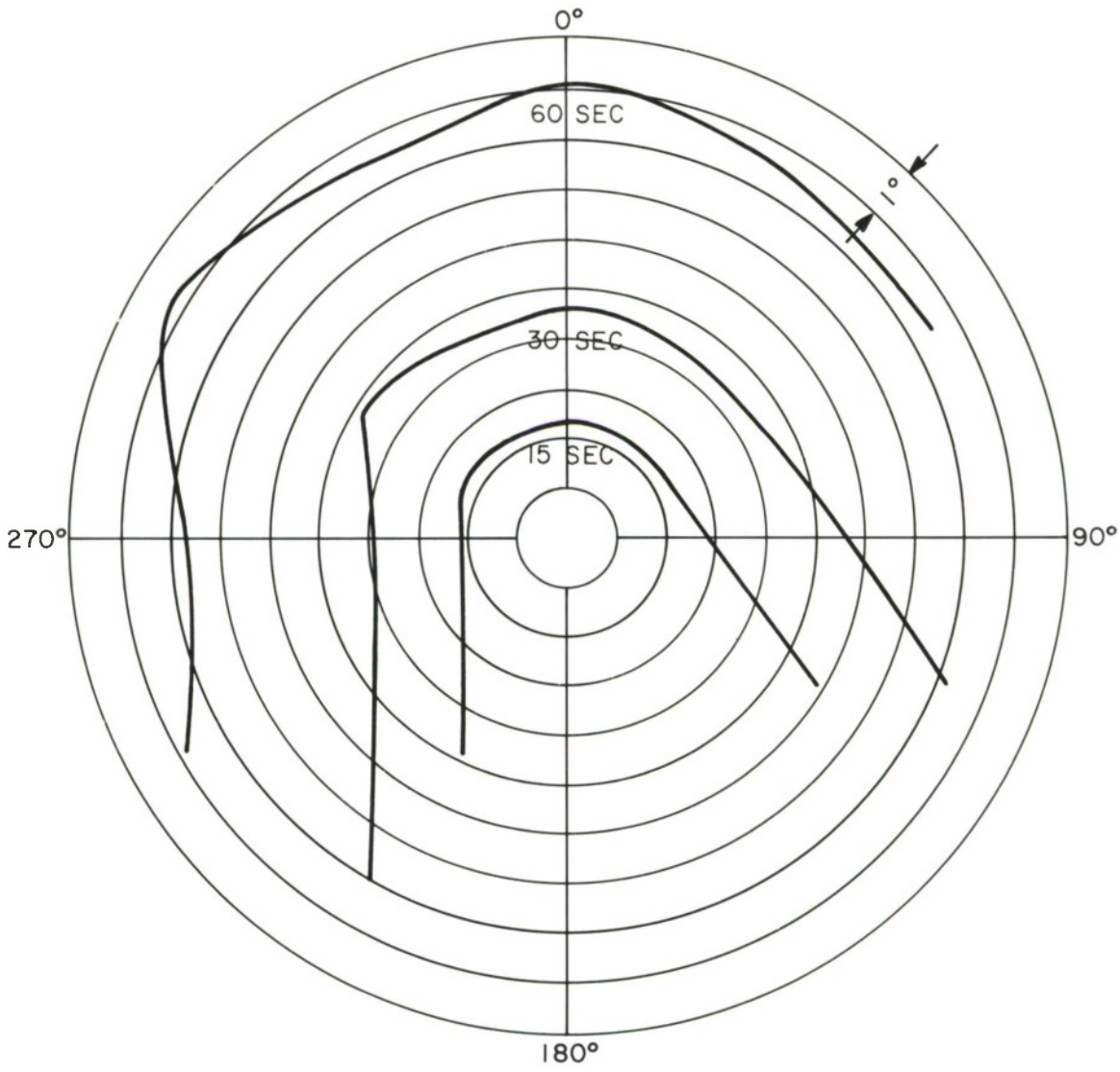


FIGURE 2.7B
23 STATION TIME DELAY CONTOURS
TWO STATION CRITERION

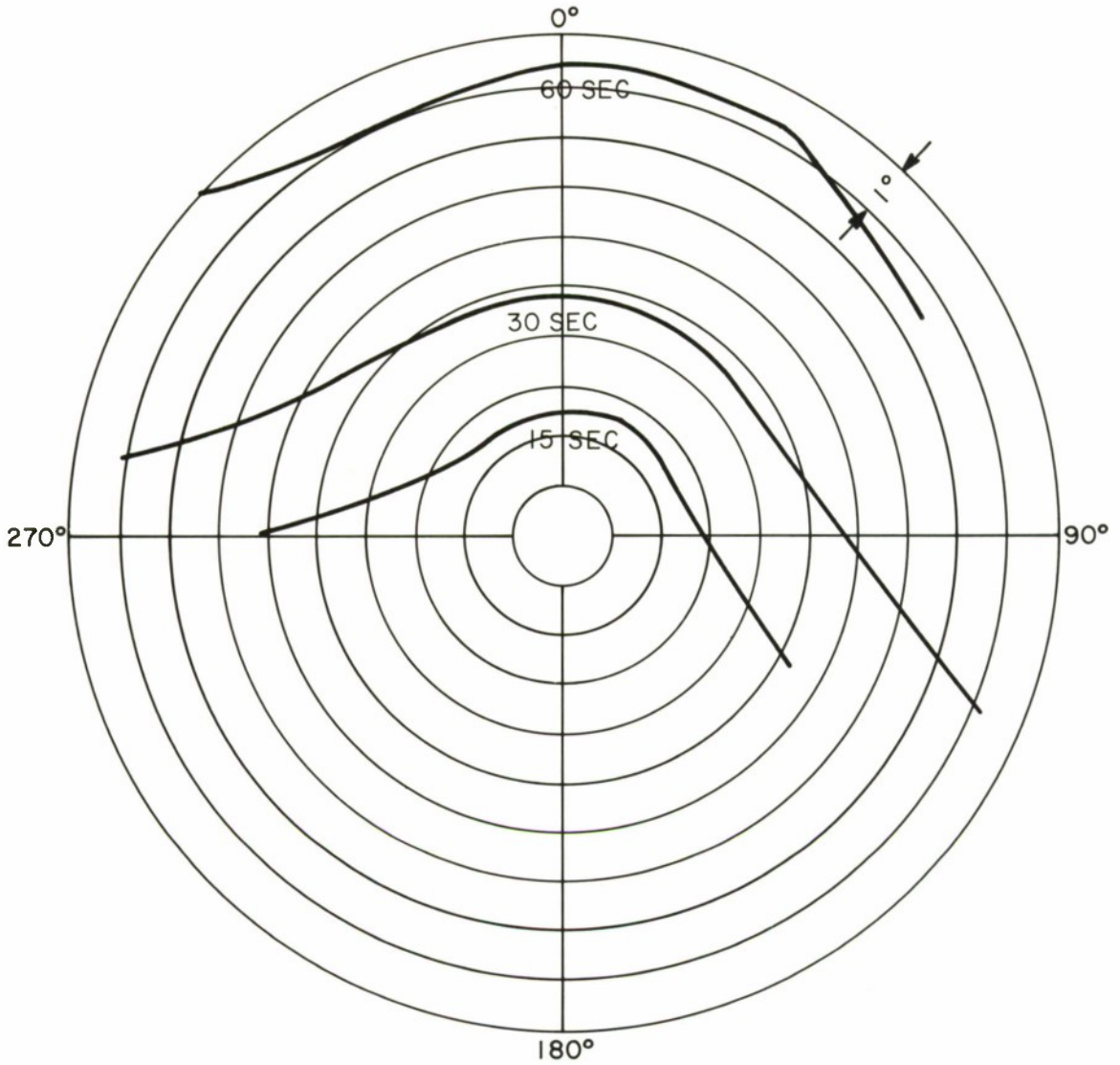


FIGURE 2.7C
23 STATION TIME DELAY CONTOURS
THREE STATION CRITERION

test would be detected. The effect of strengthening this assumption to require that at least two or at least three seismometers receive the signal from the explosion before that of the earthquake has been considered in this report. It has been demonstrated that with the set of stations originally considered, or with another set of already existing stations totaling about twenty, strengthening this assumption significantly enlarges the size of the geographic area contained within the minimum delay time contours. It would seem from a station location point of view, however, that it would be possible to operate a single station network that performed adequately. This point is illustrated by calculations based on a rather large number (116) of WWSSS stations for which the minimum delay time contours behave quite well. The conclusion suggested is that at least the locations of the seismometers are adequate for the role of the single station network, but the quality of many of the individual stations would need to be substantially improved.

Finally, it may be concluded that the proposed detection network, consisting of a continental-size array plus a network of single stations at relatively close ranges and well distributed in azimuth, would severely limit the possibilities for setting off an undetected nuclear explosion in the vicinity of an earthquake. The time constraints implied by such a network are on the order of two or three minutes for an explosion of one-half to one magnitude unit smaller size than the earthquake. This would seem to be a very severe restriction on the potential test ban violator. One problem that has not been considered in this discussion is the practical problem of carrying out the large number of calculations necessary to implement this detection network. It is possible that the number of calculations might be reduced considerably without a serious degradation in sensitivity of the detection network, but this question has not been investigated.

SECTION III

METHODS FOR EPICENTER LOCATION WITH LASA

This section is devoted to the analytical and experimental study of the errors in epicenter location that result from travel-time anomalies and additive noise. The problem with which we are concerned is that of estimation. The detection of the presence of an event plays a part here only in its role of defining which portion of the seismogram is to be considered as containing signal plus noise. The detection is performed by comparing a DIMUS sum to a threshold. The probabilities of detection and false alarm for this detector are discussed in Section IV of this report.

The discussion begins with comments concerning the properties of a DIMUS detector when the signal-to-noise ratio (SNR) is very large. It continues with a consideration of the problems caused by differences in the waveshapes observed at the individual seismometers. Then, results of simulations are presented, together with appropriate analyses, which suggest that the method described as DIMUS Energy Beam Forming should be further analyzed. The encouraging results obtained and presented herein are based on too small a data sample to allow us to draw definite conclusions. In weak noise, it will be shown that the plane wave fit method is adequate and reduces the complexity of estimation. On the other hand, it falls short of the expectations provided by the beamforming algorithms whenever the noise is not weak.

3.1 BACKGROUND AND DEFINITIONS

The major portion of this section is based on Sections III and IV of our Second Quarterly Technical Report [2].

The simulation of the effect of anomalies and noise is based on twenty-one seismograms recorded at the center elements of the subarrays at LASA-Montana for the magnitude-5.8 Kazakh event of 21 November 1965. The seismograms were aligned so as

to simulate a true azimuth of due North. Then filtered noise and anomalies were added prior to the application of various algorithms used for estimation.

An automatic detector was incorporated in order to define the time intervals of interest in a realistic manner. The integration windows were two seconds in duration and centered about the detection time indicated by the DIMUS detector. This detector and various location methods are discussed in Section IV of the Second Quarterly. We repeat a perfunctory description for convenience, with the addition of DIMUS Energy Beamforming.

3.1.1 Automatic Detection of Arrival Time

The implementation is carried out with the output of the DIMUS array. A threshold of 19 is first assumed and location in time is given by the violation of the threshold. If no such crossing occurs, the threshold is successively decreased by one unit and the process repeated until a successful detection is reached.

3.1.2 Plane Wave Fit (PWF)

First, we assume a rough knowledge of the azimuth (here due North). Then, arrival times determined by the nearest zero crossings about the time of detection are taken and used to fit a least-squares line with the abscissa determined by the geographic projection of the seismometers on a line perpendicular to the assumed direction of travel of the plane wave.

3.1.3 Analog Beamforming

The array is steered through various angles and the phased sums are computed. The two methods of beamforming estimate the azimuth by considering the maximum value of the peak of the beam pattern in the method denoted as Analog Pattern-Peak Signal (AP/P) or the maximum value of the energy in a two-second interval of the same beam in the method denoted as Analog Pattern-Energy (AP/E).

3.1.4 DIMUS Beamforming

This method differs from those previously proposed. The DIMUS beamforming algorithm performs the same computations as the analog beamforming method with the replacement of the analog waveforms by their DIMUS equivalents. It is denoted by DP/E.

3.1.5 Synthetic and Actual Seismograms

We intend to differentiate between the performance of the algorithms, under the assumptions of identical signals plus independent noises, and their performance with the actual waveforms recorded. Synthetic waveforms are obtained by constructing twenty-one identical seismograms, each being the phased sum of the original ones.

3.1.6 Assumptions and Definitions

Throughout the simulations reported, a constant phase velocity of 20 km/sec is assumed and the signal-to-noise ratio (SNR) is defined as

$$\frac{\frac{1}{2}[\text{Wave Peak}]^2}{\text{Ave. Noise Power}}$$

This definition is consistent with the conventional definition of SNR as the ratio of signal energy and average noise power when sinusoidal signals are considered. The justification for its choice is merely one of convenience.

The mean, μ, reported in the Tables represents the average of the errors made by the various algorithms. The standard deviation, σ, is that under the *a priori* hypothesis of a zero mean, and thus is computed as the rms value of the errors obtained in simulation.

3.2 DIMUS SUM DETECTION WITH LARGE SNR

When the DIMUS sum is used for detection of the presence of the signal, it is assumed that over intervals during which there is no signal present, the number of occurrences of a "+1" and a "-1" on each seismogram will average out and not contribute to any detectable congregate. It becomes apparent, however, that in the twenty-one seismograms for the Kazakh event used in previously reported simulations [2], a negative precursor as illustrated in Figure 3.1A is present in all seismograms. The fact that this precursor seems to be of negligible amplitude in the analog waveform is obviated when the DIMUS beam is formed. As a result, the main signal is presumed present prior to its principal motion and all time windows or vernier measurements used in the epicenter location algorithms are subjected to an error, as indicated in Figure 3.1B.

This artifact is responsible for the deviations from theoretical predictions and the nonzero mean of the estimate reported earlier for estimated azimuth in the presence of anomalies.

Three methods to relieve this difficulty are presented below with their merits or shortcomings. The first two are expedient measures and do not purport to be suggestions for an operational system.

3.2.1 Null-Zone Randomization

The analog waveform is first operated on by a hard-limiter with a null zone. The output of this operation is a three-level quantization representation of the analog waveform. Such precursors as mentioned above, or noise of small amplitude, would therefore be represented as zeroes rather than by their polarity. On the other hand, in order to preserve the economy of one bit of information per sample provided by DIMUS, the null zone values are randomized to plus and minus one equiprobably. The degradation of this latter step is minimized by the fact

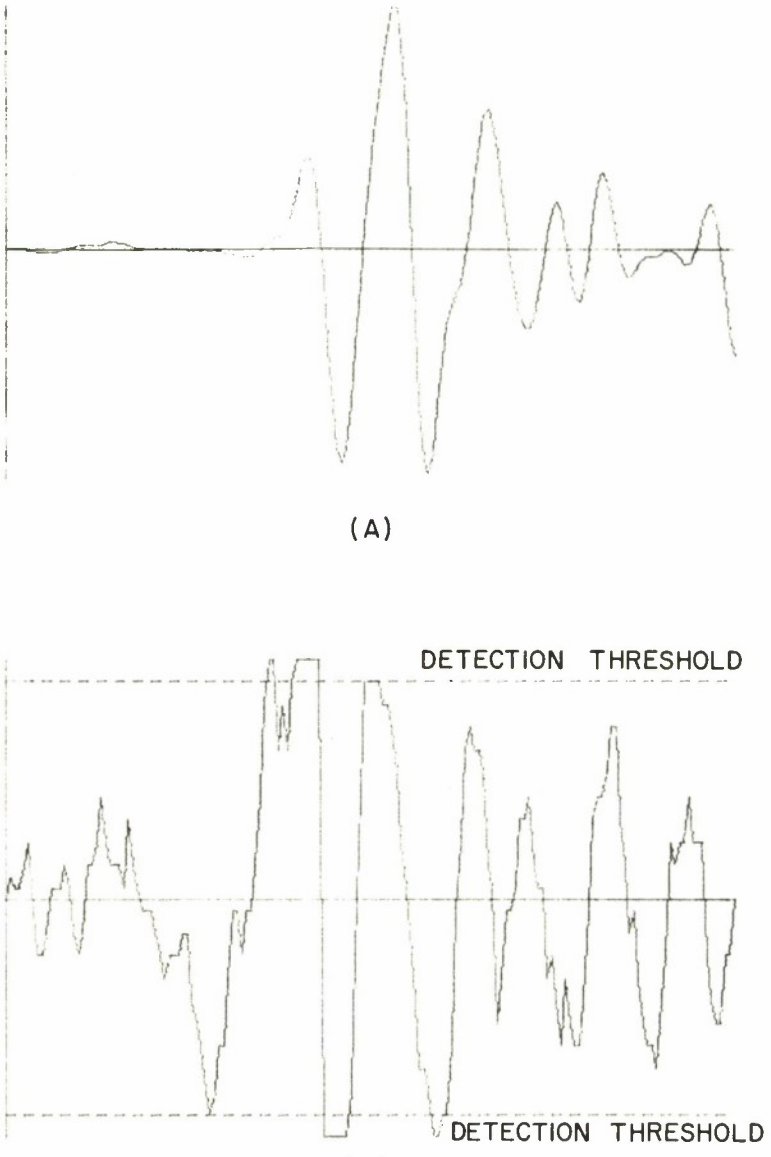


FIGURE 3.1
TYPICAL SEISMOGRAM AND DIMUS SUM
FOR STRONG EVENT

that the random noise introduced is out-of-band with respect to the signal.

Simulation results in the presence of anomalies for the same event (Kazakh, 21 November 1965) are presented in Table I. They are seen to agree with theoretical predictions derived in the Second Quarterly and suggest that the earlier difficulties were due to the effects of "false triggering" by the automatic detector.

The assignment of the null zone threshold level is a potential problem in the use of this method. It may not be set too low with respect to the ambient noise, for this would offer no help to the problem in the case of large events. On the other hand, if it is set higher than the ambient noise, small events remain in the null zone for their duration and become undetectable. The answer may be *a posteriori* setting, i.e., a level determined by the average energy received. If a small energy is detected, with or without signal present, the null zone remains negligible. If a large energy is detected on a seismogram, the relatively higher threshold would eliminate the precursor. This method, however, introduces the complexity of energy detection and memory since DIMUS can only be provided after the threshold is established.

3.2.2 Synthetic Noise

Another method of alleviating the problems caused by the small precursor consists of adding low level out-of-band noise to the analog waveform prior to hardlimiting. This would have the same effect on the precursor for DIMUS as the null zone randomization. For strong events, the out-of-band noise would be annihilated by hardlimiting when the signal is present.

Whenever the signal is small, the out-of-band noise would be eliminated by the DIMUS sum only if the number of seismograms were large. On the other hand, if the scheme resolves the difficulty of the precursor, then signals of lower amplitude than the precursor become *de facto* undetectable.

TABLE I

SAMPLE MEANS AND STANDARD DEVIATIONS VS. TIME ANOMALIES

RUN 1: 20 CASES - ACTUAL SEISMOGRAMS, ±1 SAMPLE POINT ANOMALIES

	μ	σ
PWF	0.03°	0.19°
AP/E	0.05	0.30
AP/P	0.07	0.23
DP/E	0.09	0.23

RUN 2: 20 CASES - SYNTHETIC SEISMOGRAMS, ±1 SAMPLE POINT ANOMALIES

	μ	σ
PWF	0.03°	0.19°
AP/E	0.008	0.17
AP/P	0.10	0.30
DP/E	0.008	0.27

RUN 3: 11 CASES - SYNTHETIC SEISMOGRAMS, ±1 SAMPLE POINT

	μ	σ
PWF	0.04°	0.17°
AP/E	0.06	0.11
AP/P	0.14	0.28

RUN 4: 24 CASES - ACTUAL SEISMOGRAMS, ±1 SAMPLE POINT

	μ	σ
PWF	-0.02°	0.18°
AP/E	0.08	0.34
AP/P	0.10	0.28

RUN 5: 24 CASES - ACTUAL SEISMOGRAMS, ±2 SAMPLE POINTS

	μ	σ
PWF	-0.17°	0.39°

RUN 6: 24 CASES - ACTUAL SEISMOGRAMS - NO NULL ZONE RANDOMIZATION, ±1 SAMPLE POINT

	μ	σ
PWF	0.28°	0.42°

RUN 7: 11 CASES - ACTUAL SEISMOGRAMS - NO NULL ZONE RANDOMIZATION, ±2 SAMPLE POINTS

	μ	σ
PWF	0.49°	0.70°

3.2.3 Analog Waveform Post-Decision

If we assume that at least one analog waveform is available to the processor, in addition to the DIMUS waveforms, then a comparison with the analog seismograms, post-detection, would ascertain whether or not the response of the detector is due to a precursor. Since the dilemma is present only at high SNR, there is no ambiguity in this post-detection decision. Whenever the analog waveform indicates that the DIMUS detector "triggered" on a small precursor, the indicated arrival time of the event could be modified appropriately.

3.3 WAVE SHAPE VARIATIONS ACROSS THE ARRAY

In any of the methods used in estimating the azimuth, plane wave fit using a zero crossing measure of arrival time, plane wave fit using a correlation time pick, analog pattern with energy, analog pattern with phased sum peak, DIMUS energy pattern, the performance is degraded by the fact that the signal portions of the seismograms are not identical. Indeed, it is found that, for Kazakh events, the LASA-Montana center elements B-3 and D-1, for instance, yield a disparity of one-third of a second in the time of arrival obtained by zero crossing and correlation time pick, the delay being observed in the latter method.

At the price of a sizable increase of complexity, pre-processing of the individual seismometers may be attempted. Let us assume, as a perfunctory examination of the strong event studied seems to suggest, that the seismograms are made up of the superposition of several delayed arrivals of a basic waveform. If we further assume that the delays, following the first occurrence, are not coherent, then the analog phased sum represents a first-order approximation of the basic waveform. Attempts to use this in order to remove the interference in the presence of noise introduces errors which often fail to remove the problem and lead to a performance of the epicenter location algorithm which is no better than the simpler decision to disregard

the delinquent seismograms entirely. In order to recognize this delinquency, one may either rely on the fact that this behavior seems to be consistent with the area under observation or merely use two different methods of time-of-arrival pick and set a criterion on the disparity between them for the exclusion of the seismogram.

The effect of these differences is shown in Table II where it may be observed in five of six cases that the rms errors (σ in the table) were worse when the actual waveforms were used than when the same waveform was used at each element.

3.4 COMPARISON OF PLANE WAVE FIT AND BEAMFORMING

Let us consider the idealized situation where

i) The signal, if it exists, is

$$\phi(t-t') = \sqrt{2WE_S}\,\frac{\sin 2\pi W(t-t')}{2\pi W(t-t')} \tag{3.1}$$

where t' is the time of arrival.

ii) The signal being bandlimited to (-W,W), we assume the noise to be bandlimited to the same interval with double-sided power spectrum $N_0/2$. Let the statistics of the noise be Gaussian with zero mean.

3.4.1 Weak Noise Errors for Time Picks

If a plane wave fit is performed, using the time picks for arrival of the signals with twenty-one seismometers, then in view of Equation (A-8) of Reference [2], where now the zero crossing occurs at t=1/2W, the mean square error in the time picks is

$$\overline{\epsilon_z^2} = \frac{\mathrm{var}[n(\tau)]}{[s'(\frac{1}{2W})]^2} \tag{3.2}$$

since 1/2W is the location of a zero crossing.

$$\overline{\epsilon_z^2} = \frac{WN_0}{8W^3E_S} = \frac{N_0}{8W^2E_S} \tag{3.3}$$

for a zero crossing time pick, and from Equation (A-7)

TABLE II

RUN 1 - 10 dB, 20 CASES

	ACTUAL SEISMOGRAMS		SYNTHETIC SEISMOGRAMS	
	μ	σ	μ	σ
PWF	-0.29°	0.50°	0.02°	0.45°
AP/E	-0.09	0.41	-0.07	0.32
AP/P	-0.08	0.35	-0.02	0.28

RUN 2 - 3 dB, 20 CASES

	ACTUAL SEISMOGRAMS		SYNTHETIC SEISMOGRAMS	
	μ	σ	μ	σ
PWF	-0.60	0.83	-0.08	0.46
AP/E	0.02	0.34	0.001	0.36
AP/P	-0.02	0.40	0.002	0.24

of Appendix A

$$\overline{\varepsilon_c^2} = \frac{12}{\pi^2} \frac{1}{(4W)^2} \frac{N_0}{2E_S}$$

$$= \frac{3}{8\pi^2} \frac{1}{W^2} \frac{N_0}{E_S} \qquad (3.4)$$

for a correlation time pick.

Thus, if we define G as the power gain in dB of the correlation time pick over the zero crossing time pick, we obtain

$$G = 10 \log_{10} \frac{\pi^2}{3} \approx 5 \text{ dB} \qquad (3.5)$$

3.4.2 Strong Noise Errors for Time Picks

Because of its apparent superiority, we examine the large noise behavior of the correlation time pick alone. It seems intuitively obvious that in strong additive noise, the zero crossing time pick will degenerate far sooner than the algorithm performing correlation with a replica of the signal (or an estimate of this replica).

Consider Figure 3.2 which depicts the output of the correlator used for estimating the time of arrival. Let t_0 be the true arrival time and $\hat{t}$ the estimate.

In order to calculate an approximation to the probability of such an anomalous error, let us focus our attention on a finite set of instants $\{t_i\}$ where i is an integer, defined as follows:

$$t_i = t_0 + \frac{i}{2W} \qquad (3.6)$$

$$i = -K, \ldots, +K$$

The reason for a finite limit on i is that it is unlikely that the seismograms will be used for all time in order to obtain an estimate. In other words, there is *a priori* knowledge that if an event has been detected, the signal is assumed to be in a bounded time interval. Thus we consider only 2K instants other than t_0. Typically, K will be a small integer.

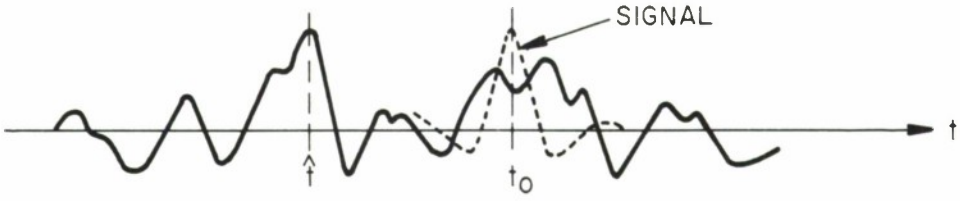


FIGURE 3.2
ANOMALOUS SIGNAL DETECTION

For any pair of instants in the interval, the waveforms have the property

$$\int_{-\infty}^{\infty} \psi(t-t_i)\psi(t-t_j)dt = \delta_{ij}E_S \tag{3.7}$$

where δ_{ij} is the Kroeneker delta, the set representing a set of sampling functions. Thus, the probability that the noisy function f(t) will peak at one of the incorrect time instants, or in its neighborhood, in the set $\{t_i\}$ is the probability of error in the communication of one of (2K+1) equally likely messages when using equal-energy orthogonal signals [3, Ch. VIII]. Let us denote by Γ the event

$$\Gamma \triangleq [f(t_i)>f(t_0) \text{ for at least one } t_i \text{ in } \{t_i\}] \tag{3.8}$$

Then we have [3]

$$\underline{P}\{\Gamma\} \simeq 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(\gamma-\sqrt{E_S})^2}{N_0}} [1-Q(\frac{\gamma}{\sqrt{N_0/2}})]^{2K} d\gamma \tag{3.9}$$

where

$$Q(\alpha) = \int_{\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\beta^2/2} d\beta \tag{3.10}$$

Approximating the Q function, we obtain

$$\underline{P}\{\Gamma\} \leq \frac{2K}{\sqrt{2\pi(E_S/N_0)}} \exp\{-\frac{E_S}{2N_0}\} \tag{3.11}$$

Since the probability that the noise causes f(t) to peak near t_i is the same as the probability that f(t) peaks near t_j for $t_i \neq t_j$, the error will be approximately uniformly distributed on (-K/2W, K/2W). Thus

$$\overline{\varepsilon^2} \leq \frac{K^4}{6W^2} \frac{1}{\sqrt{2\pi(E_S/N_0)}} e^{-E_S/2N_0} \tag{3.12}$$

where 1/2W represents the interval between two adjacent instants t_i and t_{i+1}.

For very low SNR, another approximation to the Q function yields

$$\overline{\varepsilon^2} \leq \frac{K^4}{6W^3}\left[\frac{1}{2} - \sqrt{E_S/N_0}\right] \tag{3.13}$$

Thus, if we let $\overline{\varepsilon_c^2}$ be the mean square error when the bound on the deviation is one pulse width (K=2) on either side and the correlation time pick is used,

$$\overline{\varepsilon_c^2} \leq \frac{3}{8\pi^2}\frac{1}{W^2}\frac{N_0}{E_S} + \frac{8}{3W^3}\frac{1}{\sqrt{2\pi(E_S/N_0)}}\, e^{-E_S/N_0} \tag{3.14}$$

3.4.3 Energy Beamforming

In the Second Quarterly it was shown (Appendix B) that, for weak noise, correlation time picks followed by a plane wave fit and energy beamforming are identical. The behavior of the two methods diverge in stronger noise. Apart from the requirement of having a replica of the signal in order to perform a correlation time pick, the superiority of beamforming is borne out by the tightness of the linear approximation in Equations (A-4) and (B-4) in Reference [2].

The linearization of Equation (A-4) for the time pick is valid if the estimate $\hat{\tau}$ of the time of arrival is close to the true value τ, i.e., if the variance $\overline{(\hat{\tau}-\tau)^2}$ is small. The linearization of Equation (B-4) is valid if the variance of the final estimate $\hat{m}$ of m is small.

We must bear in mind, however, that the relation between the two variances is that of the noise reduction provided by plane wave fit, i.e.,

$$\frac{\overline{(\hat{m}-m)^2}}{\overline{(\tau-\hat{\tau})^2}} = \frac{12}{N(N^2-1)} \tag{3.15}$$

Thus, the signal-to-noise ratios where either of the analyses break down bear the same relationship.

If the case for energy beamforming based on the analog waveforms is well established, that of energy beamforming based on the DIMUS waveforms remains unsettled. Table III offers a comparison based on simulations for several values of SNR. It seems at first glance that the analog and DIMUS beamforming perform with little statistical difference, at least in the limited amount of simulation presented. We now offer a rough argument for this similarity.

3.5 COMPARISON OF ANALOG AND DIMUS BEAMFORMING

We present the argument for the similarity in performance of these methods of estimating the epicenter location at the extremes of SNR.

3.5.1 Strong Noise Case

In strong noise, the property of DIMUS phased sums is that they yield a statistically consistent representation (within a constant) of the analog phased sums of the waveforms on which they are based. Thus the performance of DIMUS beamforming is expected to be comparable to analog (within the loss of $\pi/2$(2 dB) in SNR associated with DIMUS sums), for the simple reason that a similar answer is expected in each case.

3.5.2 Weak Noise Case

When the SNR is large, the polarity of the DIMUS waveform obtained from a particular seismogram, and that obtained from the noise-free signal differ only in the close neighborhood of the zero crossings of the analog signal. This deviation, furthermore, is a function of the slope of the signal at those zero-crossings.

DIMUS beamforming then maximizes,all else being assumed equal, the average pair-wise correlation of the DIMUS waveforms near the zero crossings of the analog waveforms. The number of zero crossings involved depends on the width of the integration window used in determining the energy in the phased sum.

TABLE III

COMPARISON OF ANALOG AND DIMUS BEAMFORMING

RUN 1 - 18 CASES, NO ANOMALIES

	ANALOG		DIMUS	
	μ	σ	μ	σ
10 dB	0.016	0.286	0.035	0.241
3 dB	0.081	0.339	0.068	0.340
-3 dB	0.065	0.317	0.012	0.380

RUN 2 - 20 CASES, NO NOISE, ± 1 SAMPLE POINT ANOMALIES

	ANALOG		DIMUS	
	μ	σ	μ	σ
Actual Seismograms	-0.058	0.302	-0.093	0.239
Synthetic Seismograms	-0.008	0.177	-0.008	0.277

This operation, however, is akin to minimizing the deviation from a straight line of the position of these zero-crossings or equivalently the sign changes of the DIMUS waveforms, thereby maximizing the number of coherent points. The similarity with plane-wave fit using correlation time picks is now placed in evidence.

The argument, however, requires a final comment. The correlation time pick here would be between a DIMUS representation of the seismogram and the DIMUS representation of the noise-free signal. If, as assumed, the noise is small, the identical correlation operation with analog waveforms tests their coherence of polarity. Bringing this comment to its conclusion, we expect now similarity of answer between analog and DIMUS beamforming in each particular test case and, therefore, a similarity in statistical performance.

We now must reiterate our description of the argument as being rough. We do, however, present in Table IV the detailed results of simulation and, leaving the question open, suggest that some effort be spent to further analyze the performance of DIMUS beamforming in view of its benefits in reduction of computational complexity.

3.6 SUMMARY AND CONCLUSIONS

The behavior of the error in the presence of weak noise with anomalies, presented in Table I, in conjunction with similar results in Section IV of the Second Quarterly [2] show that the accuracy of the location of the time interval of interest in the seismograms is of major importance. The solution adopted in the course of the simulations was that of null zone randomization. This proves to be useful if there is *a priori* knowledge that the event is of large magnitude. On the other hand, the alternative of using the method of analog waveform post-decision may be better suited for a system which must work with signals of a wide range in magnitude.

TABLE IV

ERROR VALUES FOR DIMUS AND ANALOG BEAMFORMING

10 dB		3 dB		-3 dB	
ANALOG	DIMUS	ANALOG	DIMUS	ANALOG	DIMUS
-0.325°	-0.300°	0.300	0.200	0.300	0.200
-0.050	-0.000	0.250	0.250	0.250	-0.050
-0.250	-0.175	-0.325	-0.325	0.550	0.325
0.300	0.250	-0.300	-0.300	-0.300	-0.300
0.250	0.300	0.350	0.350	-0.325	-0.350
0.250	0.075	-0.250	0.300	0.250	-0.050
0.275	0.200	0.550	0.550	0.350	0.350
-0.325	-0.325	-0.200	-0.250	0.249	-0.300
0.550	0.350	0.750	0.125	-0.175	-0.175
0.075	0.000	0.550	0.550	0.300	0.300
		-0.325	-0.325	0.075	0.125
		0.300	0.300	0.550	0.550
				-0.325	0.275
				-0.325	-0.325
				0.300	0.300
				0.275	0.350

Analysis and simulation of the several methods of locating the epicenter allow us to rank them according to their performance. In additive weak noise, although the performances are uniformly good, we expect, from worst to best, plane wave fit on zero crossing time picks, DIMUS energy beamforming, plane wave fit on correlation time picks and analog beamforming. In strong noise, plane wave fit on zero crossing, as suggested by the simulation, is the first one to degenerate. As the SNR is further decreased, plane wave fit on correlation time picks is expected to fall behind the beamforming methods.

The performance of DIMUS energy beamforming, in comparison with the analog beamforming, suggests that a large saving in computational complexity may exist without any appreciable loss in performance. The limited amount of simulation, together with the simplified argument of Section 3.5, without resolving the question, do establish that further study is indicated. However, it must be pointed out that the results of simulation with the Kazakh event mentioned earlier are in accordance with this conjecture.

SECTION IV

DETECTION PROBABILITIES FOR DIMUS ARRAYS

The detection capabilities of an array of hardlimited waveforms are considered in this section. The objective of this study is a clear understanding of the potential performance of present and future seismic arrays when operated as DIMUS arrays. Since probabilities of detection and false alarms are often difficult to evaluate exactly, upper and lower bounds, as well as the asymptotic form of these probabilities, will be presented. These results should provide a convenient means of evaluating the detection performance of a DIMUS array as a function of the signal-to-noise ratio (SNR) and the number of elements.

This section begins with a definition of the seismic context in which DIMUS is of potential interest. The DIMUS detection problem is then described mathematically and expressions are obtained for detection and false alarm probabilities. Borrowing results from coding theory, these expressions are bounded and their asymptotic behavior is examined. The corresponding analog array is then compared to the DIMUS array and it is shown that the SNR loss is at least 2 dB, and may be considerably larger. These results are illustrated in several ways, including standard P_d vs. SNR plots for various choices of false alarm rate and number of elements. Finally, a brief summary of the relationship between the DIMUS detection problem and coding theory is presented.

4.1 DEFINITIONS AND ASSUMPTIONS

The idealized seismic problem of interest is confined to a single sample point. It is assumed that prior to hardlimiting each of N elements has a "signal" component of amplitude s and a noise component, n_i, that is a zero mean Gaussian random variable. It is further assumed that the n_i are independent between elements and have the same variance, σ^2. The DIMUS array output is then given by

$$\sum_{i=1}^{N} \operatorname{sgn}(s+n_i),$$

whereas the corresponding "analog" array output would be given by

$$\sum_{i=1}^{N} (s+n_i).$$

The detection problem of interest is that of deciding whether or not the "signal" is present (i.e., $s \neq 0$). It is easy to show under a variety of assumptions that the optimum decision rule is simply to compare the array output (DIMUS or analog) to a threshold, which is determined by specifying a false alarm rate, by assuming a cost structure, or in some other way.

We are interested in evaluating the detection and false alarm probabilities as a function of the threshold value, the SNR (e.g., s^2/σ^2), and the number of elements, N. In addition we wish to compare the performance of the DIMUS and analog arrays. It turns out that there is a strong analogy between this problem and a familiar one in coding theory. This analogy yields many results for this problem (and in fact a more general class of problems) with relatively little effort.

Before describing the general assumptions on which these results are based it will be useful to define for the seismic case described above a few parameters of interest.

$$p[\operatorname{sgn}(s+n_i) = +1] = p[s+n_i > 0] = 1-q \tag{4.1}$$

Therefore,

$$q = \int_{s/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \triangleq Q(s/\sigma) \tag{4.2}$$

In the seismic case of interest we expect the signals to resemble a portion of a sine wave. Furthermore, for detection, the peak signal is of interest. For consistency with other discussions, we shall therefore define SNR as if s were the peak value of a sine wave, i.e.,

$$\text{SNR} \triangleq \frac{\frac{1}{2}s^2}{\sigma^2} \tag{4.3}$$

The following table presents some pairs of values of SNR and q that may be useful in subsequent discussions.

SNR =	3.3 dB	-0.8 dB	-4.5 dB	-8.6 dB
q =	0.02	0.1	0.2	0.3

4.1.1 Abstract Mathematical Model

The results presented below apply to a more general problem than the seismic case described above. To illustrate this point, the only assumptions that are crucial to the development will now be stated.

Based on the observation of N independent random variables, y_i, it is desired to determine whether hypothesis H_1 ["signal present"] or hypothesis H_0 ["signal not present"] applies. The y_i's are described by

$$p[y_i=+1|H_1] = 1-p[y_i=-1|H_1] = 1-q>1/2$$
$$p[y_i=+1|H_0] = p[y_i=-1|H_0] = 1/2 \qquad (4.4)$$

Based on these assumptions alone, it is easy to show that

$$\mu = \sum_1^N y_i$$

is a sufficient statistic and that the optimum decision rule under any of several assumptions is equivalent to comparing the sum to a threshold, K, and choosing H_1 if and only if the threshold is exceeded or equalled.

It is easily seen that the seismic problem described above conforms to this model, and it could be modified in several ways and still conform. For example, provided the y_i are independent and identically distributed, they would still conform to the model even if they were not obtained by hardlimiting the result of adding a Gaussian random variable to a constant. For example, considering an experiment with N independent tosses of a coin, y_i could be defined as +1 if a head occurs on the i-th toss and -1 if a tail occurs, and the model would still apply. In this case, the hypothesis H_0 would correspond to the coin being "fair."

4.2 p_{fa}, p_{fd} AND THEIR ASYMPTOTIC FORMS

4.2.1 False Alarms

The probability of a false alarm when the detection rule suggested above is employed is given by

$$p_{fa} = p[\mu \geq K | H_0] \tag{4.5}$$

Denoting the number of negative values of y_i by λ, we have

$$\mu = N-2\lambda \tag{4.6}$$

and therefore

$$p_{fa} = p[\lambda \leq \frac{N-K}{2} | H_0]$$

$$= (\frac{1}{2})^N \sum_{i=0}^{(N-K)/2} \binom{N}{i} \tag{4.7}$$

where it is understood that N-K is an even number. (This assumption is for convenience only and in no way represents a restriction. If N-K is odd, $\lfloor \frac{N-K}{2} \rfloor$, the largest integer less than $\frac{N-K}{2}$ replaces the latter.) Since Equation (4.7) can be cumbersome to evaluate, especially when N is large, we shall employ upper and lower bounds that were originally developed in another context. From Equation (7.8) of Fano [4] and Equation (4.7) above, we have

$$\binom{N}{\frac{N-K}{2}} (\frac{1}{2})^N < p_{fa} < \frac{N+K}{2K} \binom{N}{\frac{N-K}{2}} (\frac{1}{2})^N \tag{4.8}$$

In order to obtain some indication of the tightness of these bounds, it should be noted that they only differ by a factor of N+K/2K. If for example, we had K = N/2, the factor is 1.5.

A convenient form for these bounds can be obtained by considering their asymptotic behavior as N increases and the threshold ratio, r, defined by

$$r = K/N \tag{4.9}$$

remains constant. Defining

$$E_{fa} = \lim_{N \to \infty} \frac{-\log_2 p_{fa}}{N} \tag{4.10}$$

and following Equation 7.3 of Fano [4] we have

$$E_{fa} = 1 + \frac{1-r}{2} \log_2 \frac{1-r}{2} + \frac{1+r}{2} \log_2 \frac{1+r}{2} \tag{4.11}$$

and

$$p_{fa} \to 2^{-NE_{fa}} \tag{4.12}$$

A specific example may help to illustrate the tightness of these bounds. Assume N=21 and K=17. Then, from Equation (4.8)

$$1.1 \cdot 10^{-4} < p_{fa} < 1.23 \cdot 10^{-4} \tag{4.13}$$

Substituting r = 17/21 into Equation (4.11) yields

$$E_{fa} = 0.546 \tag{4.14}$$

and therefore from Equation (4.12)

$$p_{fa} \to 3 \cdot 10^{-4} \tag{4.15}$$

which is within a factor of 3 of the correct answer, $p_{fa} = 1.11 \cdot 10^{-4}$.

4.2.2 False Dismissals

The probability of a false dismissal when the signal is present is given by

$$\begin{aligned} p_{fd} &= p[\mu < K | H_1] \\ &= p[\lambda > \tfrac{N-K}{2} | H_1] \\ &= \sum_{i=\frac{N-K}{2}+1}^{N} \binom{N}{i} (1-q)^{N-i} q^i \end{aligned} \tag{4.16}$$

where, as before, λ is the number of negative y_i's and N-K is assumed to be divisible by 2. Returning again to Fano [4], Equation 7.18, we have

$$\frac{q(\frac{1+r}{2})}{(1-q)(\frac{1-r}{2} + \frac{1}{N})} B < p_{fd} < \frac{q(\frac{1+r}{2})}{\frac{1-r}{2} - q} B \quad \text{for } r<1-2q \tag{4.17}$$

where

$$B = \binom{N}{\frac{N-K}{2}}(1-q)^{\frac{N+K}{2}} q^{\frac{N-K}{2}} \tag{4.18}$$

As in the case of false alarms, we may define

$$E_{fd} = \lim_{N\to\infty} - \frac{\log_2 p_{fd}}{N} \tag{4.19}$$

which, following Equations 7.44 and 7.58 of Fano [4], may be written as

$$E_{fd} = \frac{1-r}{2} \log_2 \frac{1-r}{2q} + \frac{1+r}{2} \log_2 \frac{1+r}{2(1-q)} \text{ for } r<1-2q \tag{4.20}$$

Finally, we have

$$p_{fd} \to 2^{-NE_{fd}} \tag{4.21}$$

4.2.3 Tradeoffs Between p_{fa} and p_{fd}

It is customary in detection problems to illustrate possible tradoffs between p_{fa} and p_{fd} by plotting receiver operating characteristic (ROC) curves. In these plots the detection probability $(1-p_{fd})$ is shown as a function of p_{fa}, with the SNR as a parameter. For large N, the exact expressions for p_{fa} and p_{fd} (Equations 4.7 and 4.16) are cumbersome to calculate, and, furthermore, all of the calculations must be repeated if a new value of N is considered. By employing the exponents E_{fa} and E_{fd}, introduced above, most of this effort can be avoided, and satisfactory estimates still obtained.

In Figure 4.1 a family of ROC-like curves for E_{fd} and E_{fa} are presented. Each curve is obtained by varying the threshold parameter, r, while holding q fixed at the indicated value. (The SNR's that correspond to these q's in the case of signal plus Gaussian noise appear in the table immediately preceding Section 4.1.1.)

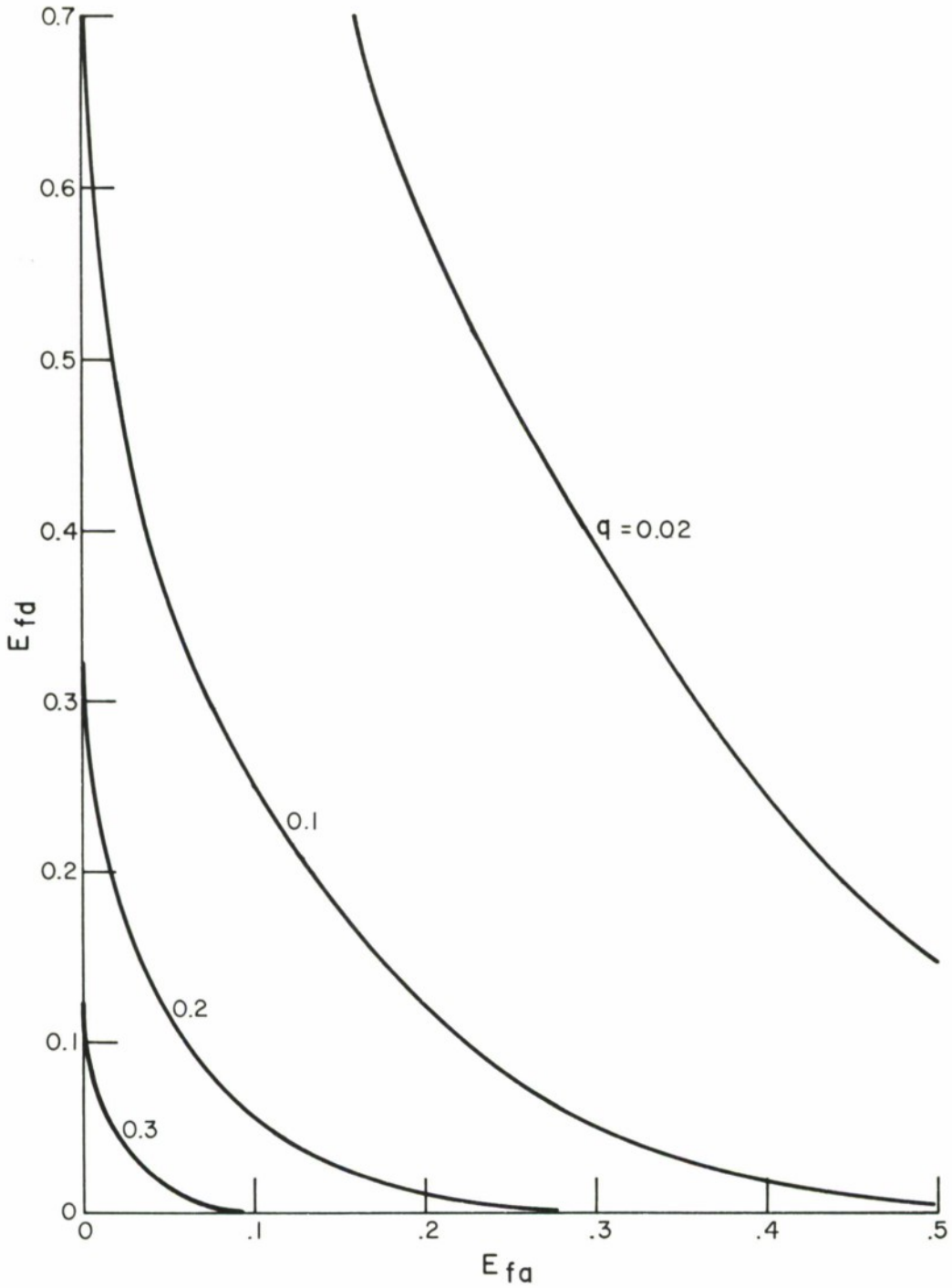


FIGURE 4.1
ASYMPTOTIC DETECTION CHARACTERISTICS

4.3 COMPARISON OF DIMUS AND ANALOG ARRAYS IN ADDITIVE GAUSSIAN NOISE

4.3.1 Properties of Analog Array

The effects of hardlimiting in the case of additive noise have been considered in a number of papers [see, for example, references 5,6]. The emphasis of these papers has usually been on comparing the output SNR's of the DIMUS and the corresponding analog array. Unfortunately, the results of these comparisons are not, in general, the answer to the following question: how much additional SNR at the input of the DIMUS array is necessary to achieve the same p_{fa} and p_{fd} as the corresponding analog array? It is this question that we wish to answer.

To consider this question we first need expressions for p_{fa} and p_{fd} in the analog case. These are, in fact, easier to obtain than in the DIMUS case. Making only a slight modification of previous definitions, we assume that the array output, μ, is compared to a threshold $K\sigma$. Then

$$p_{fa} = p[\sum_{i=1}^{N} n_i \geq K\sigma | H_0] = Q(\frac{K}{\sqrt{N}}) \tag{4.22}$$

Upper and lower bounds on p_{fa} can be obtained by using the following bounds on $Q(\cdot)$ given by Equation 2.121 of Reference [3].

$$\frac{1}{\sqrt{2\pi}\gamma} e^{-\gamma^2/2}(1-\frac{1}{\gamma^2}) < Q(\gamma) < \frac{1}{\sqrt{2\pi}\gamma} e^{-\gamma^2/2} \quad \gamma>0 \tag{4.23}$$

If, as before, we hold $r = K/N$ fixed and let N increase, then the argument of $Q(\cdot)$ in Equations (4.22)-(4.23) will be equal to $r\sqrt{N}$, the upper and lower bounds will have the same exponent, E_{fa}, given by

$$E_{fa} = \lim_{N\to\infty} - \frac{\log_2 Q(r\sqrt{N})}{N} = \frac{1}{2}r^2 \log_2 e \tag{4.24}$$

As before, then

$$p_{fa} \doteq 2^{-NE_{fa}} = 2^{-N[\frac{1}{2}r^2 \log_2 e]} \tag{4.25}$$

The probability of a false dismissal is given by

$$p_{fd} = p[\sum_{i=1}^{N}(s+n_i)<K\sigma] = Q(\frac{N\frac{s}{\sigma}-K}{\sqrt{N}}) \qquad (4.26)$$

In the same manner that Equation (4.24) was obtained, we may obtain

$$E_{fd} = \frac{1}{2}(\frac{s}{\sigma}-r)^2 \log_2 e, \quad s/\sigma>r \qquad (4.27)$$

Combining Equations (4.24) and (4.27) yields

$$E_{fd} = (\frac{s}{\sigma}\sqrt{(\log_2 e)/2} - \sqrt{E_{fa}})^2 \qquad (4.28)$$

The asymptotic form of p_{fd} may now be written as

$$p_{fd} \rightarrow 2^{-N[\sqrt{(s^2/2\sigma^2)(\log_2 e)}-\sqrt{E_{fa}}]^2} \qquad (4.29)$$

4.3.2 Comparison of DIMUS and Analog Arrays

Three different ways of viewing the difference between the DIMUS and analog arrays in terms of the exponents are presented in Figure 4.2-4.4. Figure 4.2 presents the ROC-like curves introduced earlier for both arrays and for SNR's of -0.8 dB and -8.6 dB. Figure 4.3 considers a fixed E_{fa} and plots SNR vs. E_{fd}. Finally, Figure 4.4 plots the additional SNR necessary for the DIMUS array to achieve the same E_{fd} as the analog array, as a function of r, or, equivalently, as a function of E_{fa}. Strictly speaking, Figure 4.4 is based on the limiting case of $E_{fd} \rightarrow 0$, and this point merits a few additional comments. Returning to Figure 4.3, it is clear that the SNR difference between the DIMUS and analog arrays is a function of E_{fd}. In general applications, however, we are principally interested in small values of E_{fd}. For example, with N = 100 elements, a p_{fd} of 0.1 implies an E_{fd} of 0.03; whereas a p_{fa} of 10^{-6} would imply an E_{fa} of 0.20. As suggested by this example we expect to be concerned with values of E_{fd} near zero, and values of E_{fa} that are significantly larger. With this in mind, we shall take as a standard of comparison the SNR difference

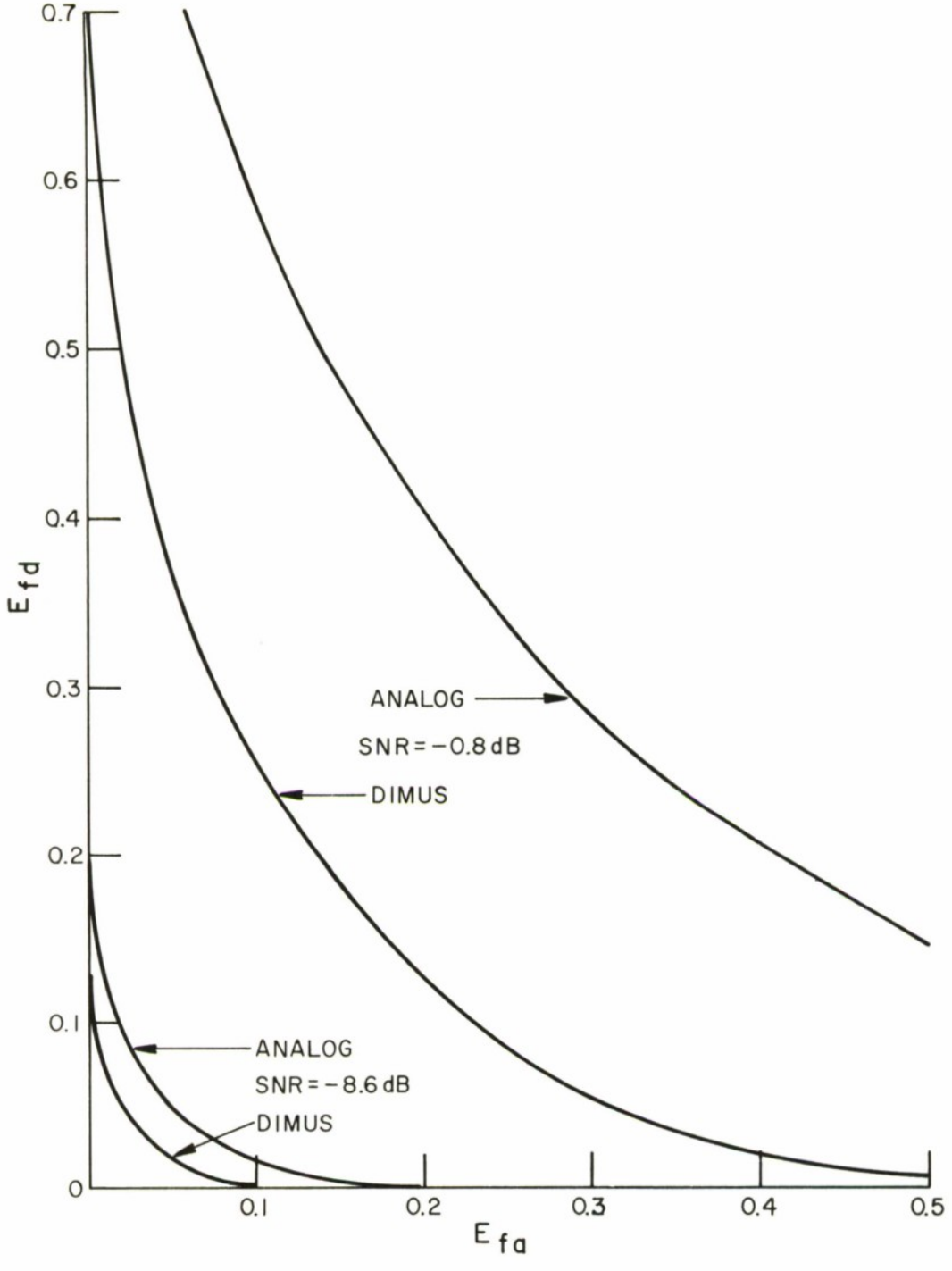


FIGURE 4.2
COMPARISON OF DETECTION CHARACTERISTICS

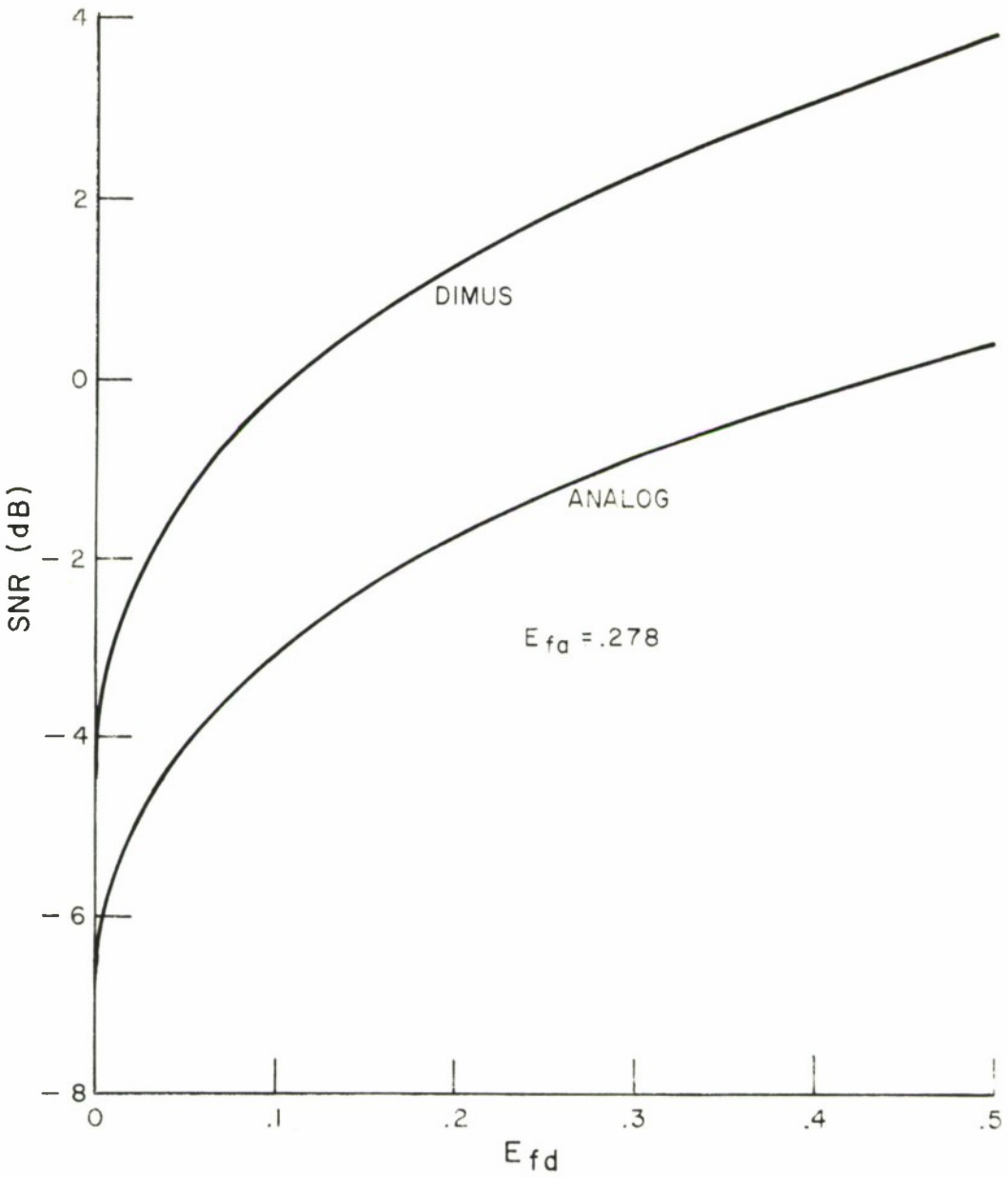


FIGURE 4.3
COMPARISON OF SNR'S FOR
E_{fa} AND E_{fd}

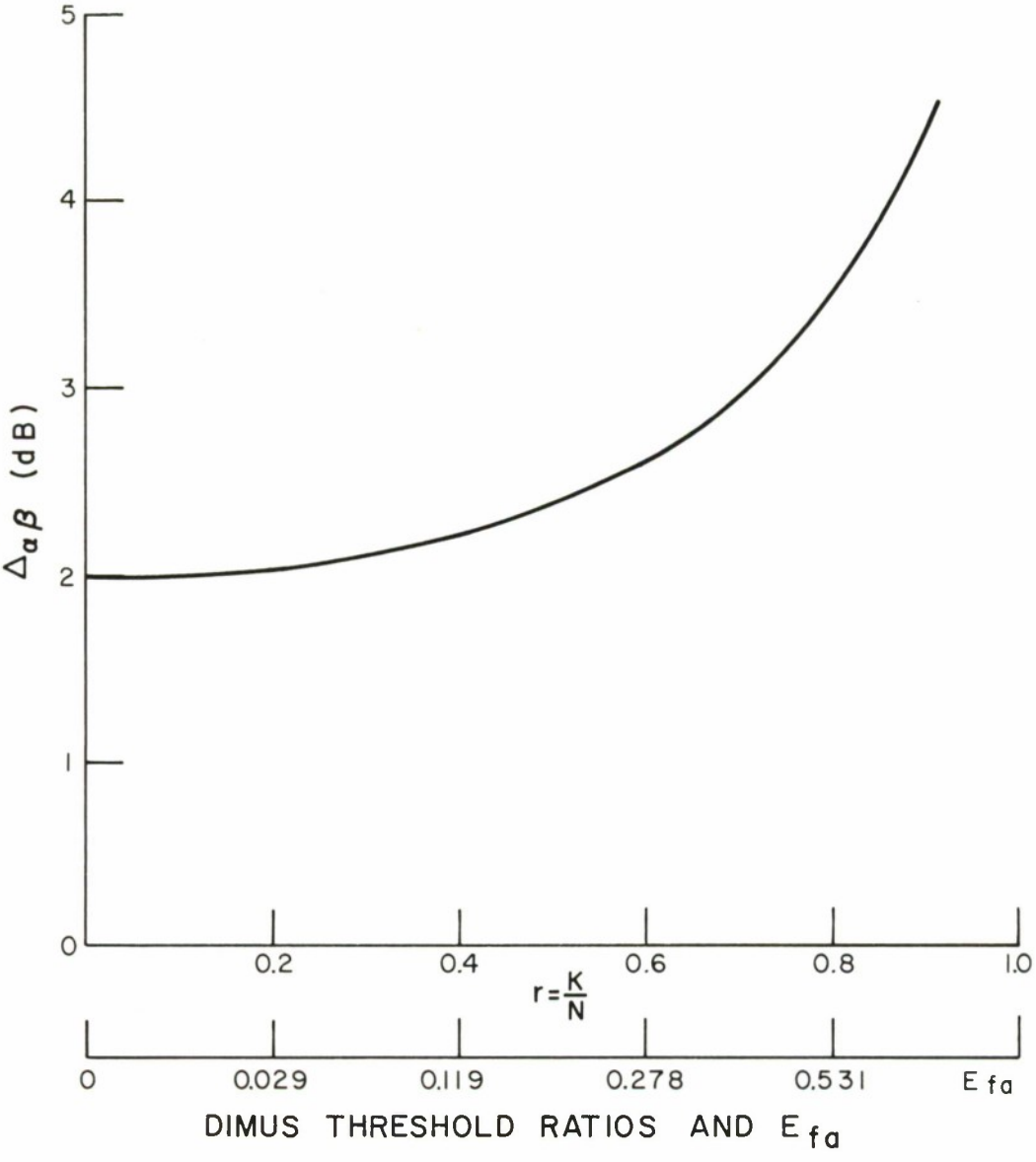


FIGURE 4.4

DIMUS SNR LOSS
VS.
THRESHOLD SETTING

as $E_{fd} \to 0$. (Figure 4.3 suggests that this will give a smaller difference than applies at any nonzero value of E_{fd}.)

The $E_{fa} = 0.278$ point of Figure 4.4 is obtained by measuring the difference (2.6 dB) between the curves of Figure 4.3 at $E_{fd} = 0$. The other points could, in principle, be obtained similarly, although in fact it is easier to develop equations specifically for this purpose. From Figure 4.4 it is clear that the SNR loss of DIMUS over the analog case increases with K from a minimum value of just under 2 dB, which is exactly the $\pi/2$ factor in power that is often quoted as the loss due to hardlimiting.

Figures 4.2-4.4 have the advantage of providing relationships that are independent of N; they have the disadvantage, however, that some additional calculations are necessary to go from the exponents (assuming a particular N) to the more familiar quantities p_{fa} and p_{fd}.

Figures 4.5 and 4.6 present $p_d = 1-p_{fd}$ vs. SNR curves for several assumed values of N, K and p_{fa}, for both the analog and DIMUS arrays. These curves indicate in a more concrete fashion some of the conclusions hinted at earlier and illustrate typical values that result when parameters that are thought to be relevant to seismic applications are employed. It should be noted that the motives for the choices of N = 21, 80 and 127 are somewhat mixed. Some experiments have previously been performed using the 21 center elements of LASA-Montana. Additional experiments are being planned that may use approximately 80 elements of LASA-Montana, chosen to minimize noise correlation between elements. Finally, we understand that LASA-Norway will eventually include in excess of 127 short period instruments. N=127 is the largest number that could be quickly and easily handled with the computer and subroutines available for these illustrative calculations.

The number"2 dB"is often quoted as the degradation resulting from bandlimiting. Figure 4.4 suggests that the

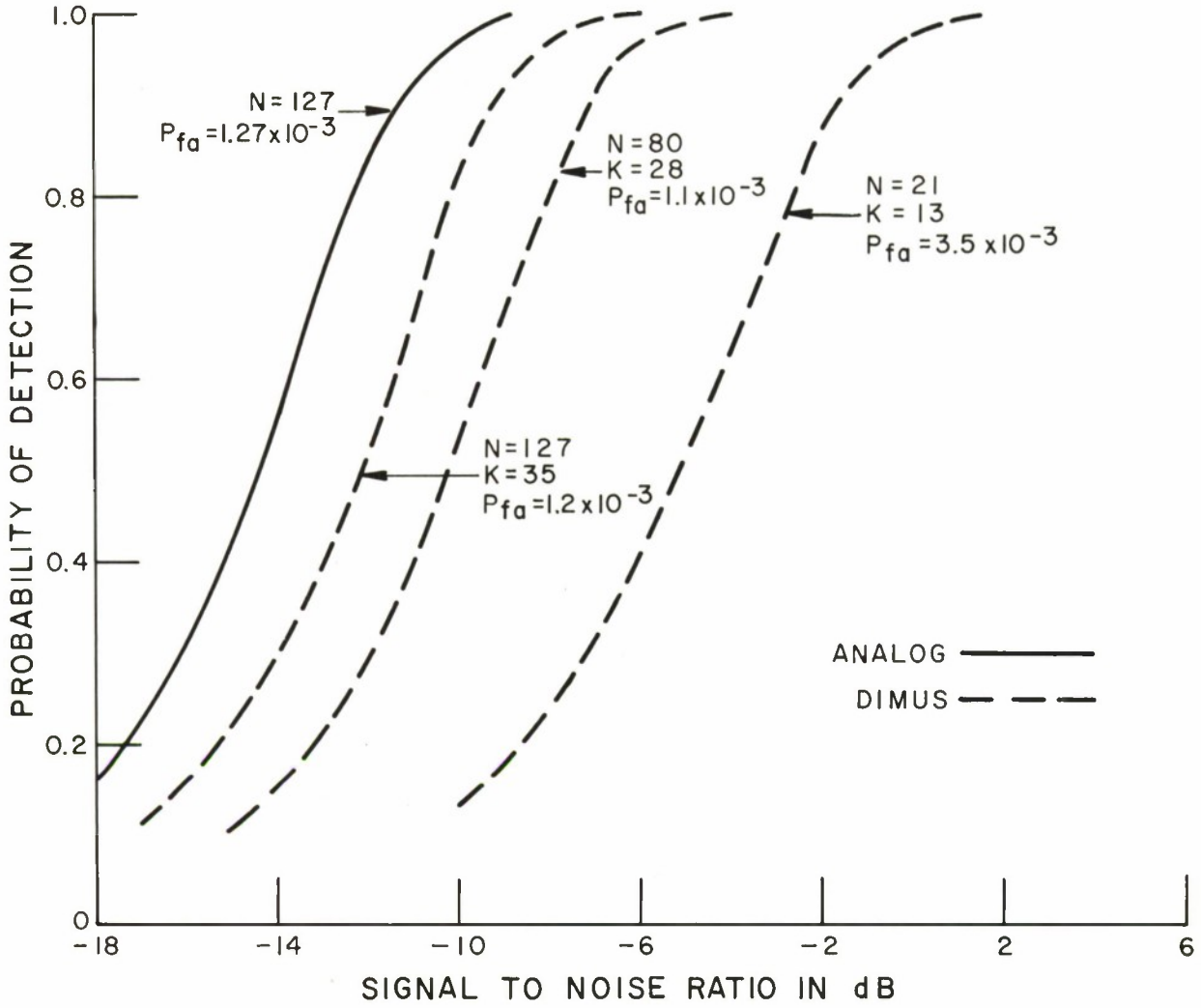


FIGURE 4.5
PROBABILITY OF DETECTION CURVES
($P_{fa} \approx 10^{-3}$)

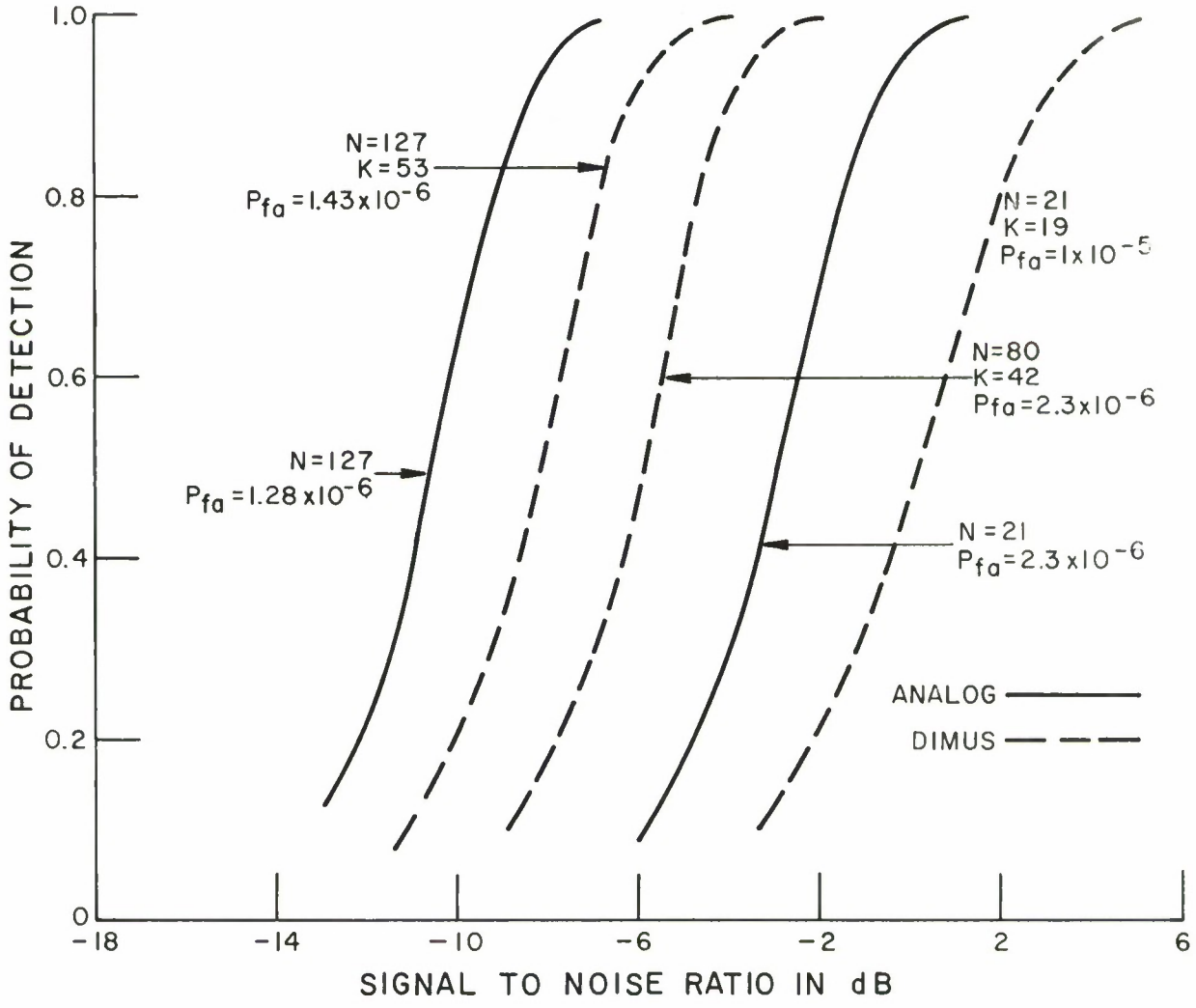


FIGURE 4.6
PROBABILITY OF DETECTION CURVES
($P_{fa} \approx 10^{-6}$)

number applies to a limited range of cases if probabilities are considered. Figure 4.6 makes this point explicit. The DIMUS and analog curves are presented for the N=21 case. For DIMUS the false alarm probability is 10^{-5}, for the analog array it is somewhat lower. If the analog threshold were adjusted to yield the same p_{fa}, the curve would move left on the figure. Even in its present position, however, it is clear that there is more than a 2 dB discrepancy between the curves, especially at the higher detection probabilities.

4.4 RELATIONSHIP OF DETECTION RESULTS TO CODING THEORY RESULTS

It is no accident that the bounds provided in Reference [4] are useful in simplifying the results of Section 4.2. Those familiar with bounds on the probability of error for a binary symmetric channel (BSC) will, after some interpretation, recognize that p_{fd} is equal to the probability of error given by the "sphere packed" bound. This fundamental bound of information theory provides a lower bound on the probability of error for the binary symmetric channel. However, for the detection problem, the sphere packed bound is not a bound at all, but is exactly equal to the probability of a false dismissal.

A list of the relationships between the detection and coding parameters is given in Table V. It is significant to note that the trade-off in coding theory between the value of the sphere packed exponent and the rate of communication, is exactly analogous to the trade-off between the values of E_{fd} and E_{fa} given by Figure 4.1. The point where E_{fa} achieves its maximum value, as $E_{fd} \to 0$, is of significance in coding theory since it is equal to the channel capacity of a BSC. It is also of significance in detection theory since it corresponds to the trade-off which is of most practical interest.

The analog between detection theory and coding theory does not stop with the relationship between an array of hard-limited data and data passed through a BSC. Indeed, a similar relationship exists between the bounds for the array of analog

TABLE V

RELATIONSHIP BETWEEN DETECTION AND CODING PARAMETERS

Parameter	Detection Theory	Coding Theory
N	The number of elements in the array.	The number of letters in each code word
q	The probability of receiving a -1.	The crossover probability for the BSC.
p_{fd}	The probability of a false dismissal.	The probability of an error for the sphere packed bound
E_{fd}	The false dismissal exponent.	The sphere packing error exponent.
max E_{fd}	The maximum false dismissal exponent occurring when $E_{fa}>0$.	The zero rate sphere packing experiment
E_{fa}	The false alarm exponent.	The rate of communication.
max E_{fa}	The maximum false alarm exponent occurring when $E_{fd}\rightarrow 0$.	The channel capacity of a binary symmetric channel.

data and the results known about the infinite bandwidth Gaussian channel. It seems evident that coding theory techniques developed for more complex communications channels, those involving $K>2$ symbols, can be exploited in the treatment of more complex detection problems, and perhaps, the reverse is also true.

SECTION V

AUTOMATIC pP TEST

An automatic pP test that employs seismograms obtained from a continental-size seismic array has been reported in previous reports [1,2,7]. Recently, the properties of this test have been re-examined with the objective of answering more clearly two questions: under what circumstances does this test perform better than other methods and what components of the test are essential to its success? To consider these questions, we have reconsidered previously processed data from a variety of seismic events: earthquakes with depths ranging from 5 km to 150 km and underground nuclear explosions.

The automatic test is composed of three basic parts: (1) a procedure for aligning seismograms for coherent addition of the pP phase arrivals; (2) an energy measurement for detection of the pP phase arrival on the phased sum seismogram; and (3) a measure of signal correlation among the seismograms to distinguish the pP phase arrival, which would be fairly consistent across the array, from spurious high energy peaks on individual records. The test has been described in more detail elsewhere [2,7]. A critical examination will be given these components of the automatic test in an attempt to justify their necessity. If any is not required for good test performance, then it may be possible to simplify the automatic test significantly. This is the case for the shallow pP test which searches for event depths in the 10 to 40 km region. In this region, it is only necessary to make a gross correlation measure as defined by the coda-correlation discriminant [Section IX of Reference 17], and then, if this correlation measure indicates that the event is in the 10 to 40 km region, to calculate simple energy ratios without correlation weighting. Such a simplification is not evident for the extended test which searches for event depths down to 200 km. The value of pP phasing will be demonstrated with three moderate depth earthquakes (about 50 km).

The correlation measure alone has proved not to be a suitable statistic for detection of pP, but its value for reducing "false" energy peaks will be illustrated.

Consideration will also be given to the need for widely spaced (continental-size) seismometer arrays, rather than moderate-size arrays, such as LASA-Montana. The value of the larger arrays for reducing local noise, aftershocks, and local-to-the-seismometer crustal reverberations in the pP phased seismogram and in calculations of the η test statistic, will be noted.

Finally, in previous applications of the test, spurious peaks which tend to make confident depth estimates difficult, if not impossible, occasionally occur. For these events erroneous conclusions may result as is the case of a nuclear explosion in the Novaya Zemlya region. These extraneous peaks can be explained by a multilayered model of the earth's crust. While we know of no way to eliminate this problem, it does appear that a knowledge of the geology of the epicenter region may be very useful in interpreting the test results.

5.1 ARRAY PROCESSING FOR pP ENHANCEMENT

It is well-known that array processing can be used to enhance desired signals in the presence of undesirable "noise" when dealing with signals more strongly correlated than the surrounding noise. This is generally the case with waveforms of teleseismic events. The pP phase arrival is the signal of interest, and the surrounding energy is considered "noise." Portions of the coda, however, can be as strongly correlated as the pP signal across an array.

Earthquake aftershocks, crustal reverberations and local noise often make identification of the pP phase arrival difficult on individual seismograms. An array such as LASA-Montana can alleviate the local noise problem, but local crustal reverberations in the coda can be quite correlated. This is evidenced by experimental runs of the automatic pP test

with LASA data. When the individual seismograms' signal-to-noise ratios were weak, the case where the automatic test is of most interest, multiple peaks occurred in the test statistic [Section V of Reference 1], probably due to crustal reverberations beneath the seismometers. Larger, continental-size arrays can eliminate local (to the seismometer) crustal reverberations since the earth's crustal structure is different below each receiver. At the same time, these larger arrays help eliminate aftershocks. With large arrays, the relative P-pP delay times can vary by as much as 6 seconds for a 200 km depth event, whereas aftershocks follow main P by a constant delay time at all seismometers, provided the aftershock hypocenter is the same as the initial shock hypocenter. Hence, as signals are phased for pP by delaying them relative to P, the aftershock signals will not add coherently. Crustal reverberations local to the event are another problem, and will be discussed in a later section.

To demonstrate the power of array processing for enhancement of the pP phase arrival, and hence to show the advantage the automatic test has over techniques which utilize only individual seismograms, it was felt that clear visual evidence was required. With this goal in mind, a technique was developed to combine a set of seismograms for one event from widely spaced seismometers into one phased-sum seismogram in such a fashion that, regardless of the event's depth of focus, the pP phase arrivals on the individual records would add coherently in the sum seismogram. For a given depth of focus, the P-pP delay time increases nonlinearly with range. For a fixed range with a constant velocity assumption the P-pP delay time increases linearly with depth. In forming the sum seismogram the seismometer nearest the event is taken as a standard. Then the seismograms from the more distant stations are compressed in time by deleting points so that the P-pP delay times are comparable to the P-pP delay time at the nearest station. Since the delay times increase linearly with depth, points are removed from seismograms at equally spaced intervals and the remaining

points shifted together as if the points removed were never there. For typical arrays, this means that at most about one point in ten is removed. Hence, the compression should introduce negligible waveform distortion on the majority of the seismograms. The "compressed" seismograms are then simply summed by aligning on main P, and the pP phases should add coherently, regardless of the true depth of the event.

Three events are presented where the pP (or sP) phase cannot be located easily on individual seismograms, but can be located on the phased sum seismogram. The first event, 6 October 1962, was an Ecuador earthquake reported earlier [Section VI of Reference 2] as difficult to analyze; our tentative results indicated a 77 km depth. C&GS, however, had very good evidence from nearby S-phase arrivals that the event was about 150 km deep. Figure 5.1 shows the individual seismograms for this event, which have been filtered with a 0.6-3.25 Hz bandpass filter. It is not possible to identify with any confidence the pP or sP phase arrival on any of these seismograms even with *a priori* information that this is a 150 km event. Figure 5.3A shows the phased sum of the "compressed" versions of the seismograms shown in Figure 5.1. Here, there is clearly a phase arrival about 41.5 seconds following the P-phase arrival. The nearest station is 41.6° from the event; this corresponds well to the P-sP delay time for a 150 km earthquake. It is not unusual for the sP phase to be much stronger than the pP phase, which seems to be the case for this event.

The second event, 2 January 1964, was a magnitude-5.1 Kamchatka earthquake, reported, when only the shallow pP test existed, as deeper than 40 km [Section X of Reference 7]. The unfiltered seismograms for this event are shown in Figure 5.2. On three of the five records, the pP phase arrival is difficult to identify, or its delay time relative to P is difficult to measure. In addition to strong noise, there are strong codas on these records. Notice what happens, however, on the phased sum of the "compressed" seismograms shown in Figure 5.3B. The

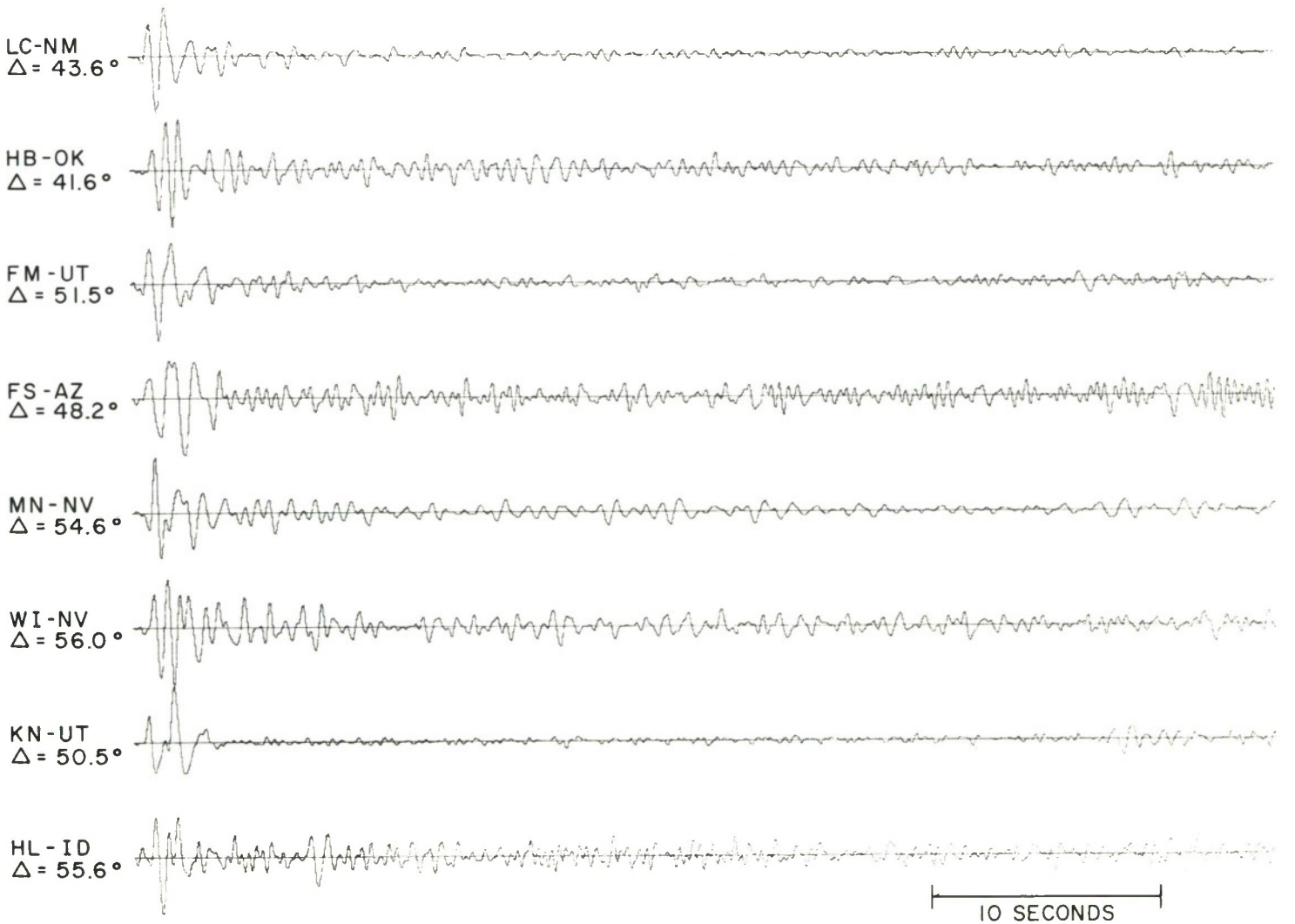


FIGURE 5.1 - SEISMOGRAMS FOR 6 OCTOBER 1962 EVENT (FILTERED)

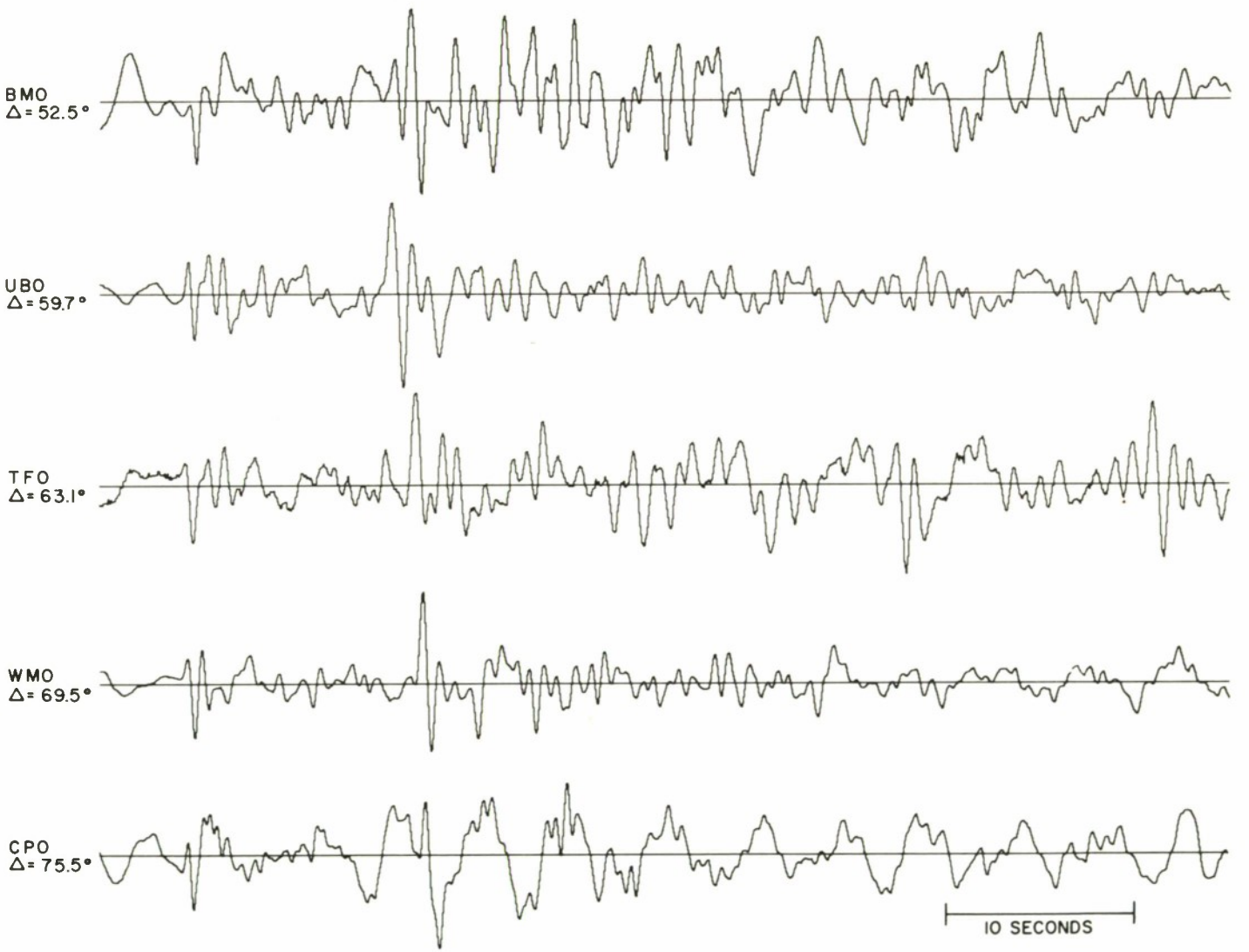


FIGURE 5.2 - SEISMOGRAMS FOR 2 JANUARY 1964 EVENT

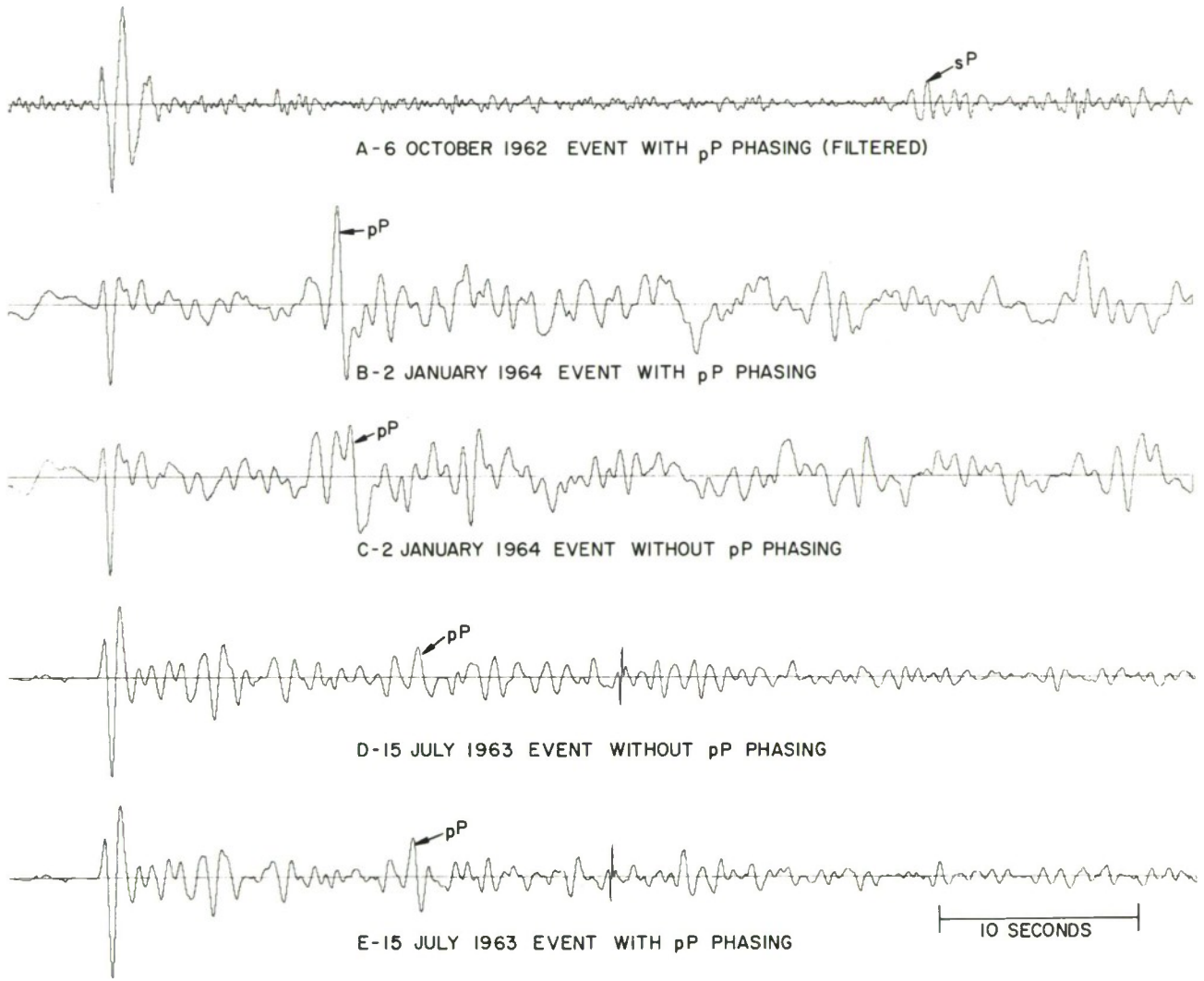


FIGURE 5.3 SUM SEISMOGRAMS

pP phase arrival is clear and the general coda level is well down relative to that of the individual seismograms. It is easy to make a confident depth estimate using this phased sum record. The measured P-pP delay time is 11.4 seconds; the nearest station range is 52.4°. This gives a depth estimate of 42 km, which compares well with the 40 km estimate of the Vela Seismological Center. Discussion of a third example, the 21 March 1962 Alaska earthquake is deferred until the next section, since it is used to illustrate several additional points, and fits more logically into that discussion.

The two events just discussed illustrate the pP or sP enhancement power of array processing designed to suppress interfering coda and noise which might be strong enough to make identification on individual seismograms difficult. A reasonable question at this point is whether or not these phase arrivals would show up just as clearly on the sum of "uncompressed" seismograms. The next section considers this question.

5.2 ALIGNMENT OF SEISMOGRAMS FOR THE pP PHASE ARRIVAL

In order to utilize the coherent property of the pP phase arrivals at the elements of a seismometer array, the individual signals must be added in phase. This means the relative P-pP delay times vs. depth must be known, or calculated, to within about 0.2 seconds for each pair of array elements. As was pointed out in an earlier report [Section VI of Reference 2], this level of precision appears feasible with the delay time calculations used by the automatic test. There are indications, however, that severe anomalies in P-pP delay times are not unusual. The Seismic Data Laboratory reports up to 0.5 second variation between the P-pP delay time indicated by JB tables and that measured at the TFO Extended Array [8]. Since the TFO Extended Array has only a 4° aperture and the automatic test used arrays with 10° to 30° apertures, delay time discrepancies could be even more severe than 0.5 second.

This seems not to be the case though, at least on a majority of the seismograms used in the following discussion.

Experiments indicate that P-pP delay times can be calculated precisely enough to enhance significantly the appearance of the pP phase arrival on a sum seismogram phased for pP as compared to a sum seismogram formed simply by adding seismograms phased on main P. Three events are presented to substantiate this result.

The sum seismogram without pP phasing for the first event, the 2 January 1964 event of the last section, is shown in Figure 5.3C. The pP phase arrival is indicated on the plot. This should be compared to Figure 5.3B, which is the sum seismogram phased for pP. The increased clarity of Figure 5.3B is evident.

The sum seismogram without pP phasing for the second event, the 15 July 1963 Kamchatka earthquake, is shown in Figure 5.3D. Below this is the sum seismogram phased for pP, Figure 5.3E. Again, the pP phase arrival is signficantly more well-defined on the sum seismogram that is phased for pP.

The third event presented illustrates several points. This is an Alaskan earthquake of 21 March 1962. Figure 5.4A shows the individual siesmograms for the event. Figure 5.4B shows the sum seismogram with phasing for pP. Figure 5.4C shows the sum seismogram without phasing for pP. The strong signal following main P by 55 seconds appears more clearly in the sum seismogram without pP phasing than in the sum seismogram with pP phasing. It seems this is an aftershock. Both we and C&GS have reported this "aftershock" as the pP phase arrival [Section VI of Reference 2]. The relative delays used to form the pP phased seismogram have decorrelated the aftershock waveforms on the individual records, hence the muddled appearance of this phase arrival in Figure 5.4B. Simultaneously, the pP phasing has enhanced the true pP phase arrival as indicated in Figure 5.4B, while this arrival is not clear in Figure 5.4C. This is an example of how a large array (1100 km

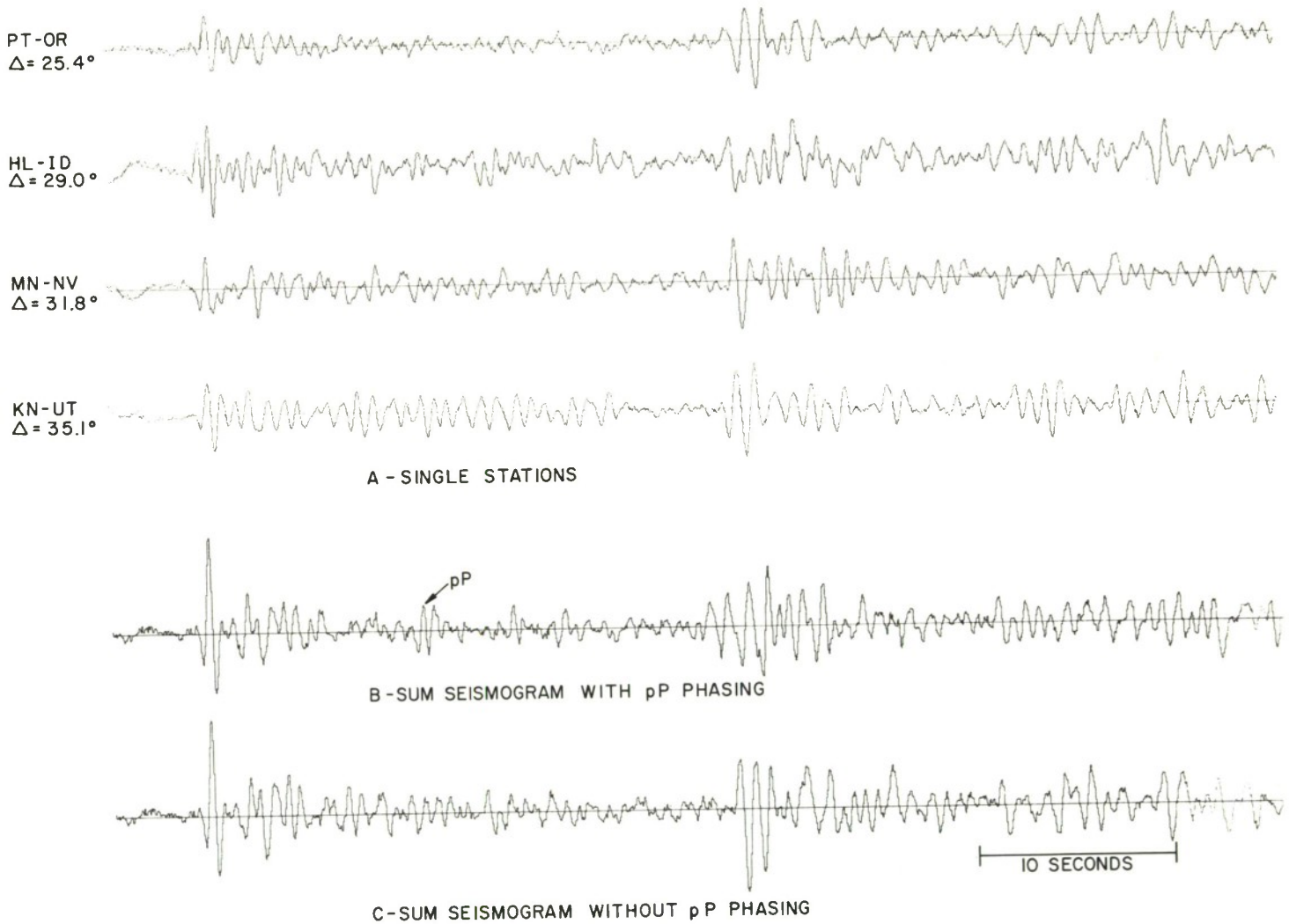


FIGURE 5.4 - SEISMOGRAMS FOR 21 MARCH 1962 EVENT

aperture) can alleviate the aftershock problem. The η statistic was larger for the pP phase arrival than for the aftershock arrival, even though the aftershock clearly shows more energy than the pP phase. This event also illustrates the point of the first section: when the pP phase cannot be identified clearly on individual seismograms, it may be distinguishable on the sum seismogram. It should be noted that this event uses only four stations, a marginal number for good automatic test performance with the strong coda of the records used. These four stations, however, have a good "average" spread in range, that is, all stations are separated by at least 300 km. Also, the array is reasonably close to the event, the station ranges varying from about 25° to 36°. In this region, the travel-time tables have their greatest slope, giving maximum relative P-pP delay times for an array with a given aperture.

5.3 ENERGY AND CORRELATION DETECTION OF pP

The development of the automatic pP test was an *ad hoc* process. The test began by measuring the energy in a one-second window at the assumed location of the pP phase arrival on the sum seismogram and comparing this to the surrounding average energy level. The test was applied only to the 10 to 40 km depth region, and peak values of the calculated energy ratio were selected as the pP phase arrivals. It was reasoned at that time that spurious high energy peaks on one seismogram could possibly change the location of the energy ratio peak from the true pP location since the summing technique tended to emphasize seismograms with high signal-to-noise ratios. Thus, the average paired correlation of the signals in the trial locations of the pP arrivals on the individual seismograms was introduced as a weighting factor, with the hope of eliminating these spurious energy peaks. In retrospect, it appears this correlation measure was not required for this purpose. In 24 earthquake trial runs, the correlation weighting factor did not once change the location of the test statistic peak, that is, the

energy ratio and the energy ratio weighted by the average paired correlation peaked at the same place for each trial run. However, in comparing energy ratio peaks of earthquakes with those of nuclear explosions, one finds roughly equivalent results. If a reasonable energy ratio threshold is determined for the 24 earthquake runs in such a fashion that peaks of the energy ratio above this threshold are accepted as valid pP phase arrivals, and peaks below the threshold value are rejected as spurious energy peaks, then for 100% detection of valid pP arrivals, 14 out of 22 runs which should not indicate a pP phase arrival in the 10 to 40 km region would have energy ratios large enough to exceed this threshold; 11 of these 14 "false" energy peaks occur with shots. If now the average paired correlation is introduced to weight the energy ratio and a new threshold similar to the above is determined for the weighted energy statistic, there are only three "false" peaks, two of them from shots, which would be accepted as valid pP phase arrivals. The one false earthquake peak was for the 15 July 1963 Kamchatka event, which will be considered along with the two shots, in the discussion of crustal reverberations.

A coda correlation discriminant was developed simultaneously with the shallow version (10 to 40 km) of the automatic pP test. This discriminant is quite effective at separating 10 to 40 km earthquakes from other events. One major simplification of the automatic pP test in this shallow region might be to run first the coda-correlation test to determine if an event is to be accepted as "in the 10 to 40 km region." If so, then the energy ratio would be calculated without correlation weighting and the peak would be selected as the pP location. If the event is not "in the 10 to 40 km region," according to the coda correlation test, then no pP test need be run. With this scheme, only one shot and one earthquake produced "false" results.

Since these results were obtained, the pP test has been extended to search for possible event depths down to 200 km. A measure similar to the original coda correlation discriminant has not yet been developed for deeper events, thus the pP test still has some significant value for discrimination purposes. For deeper test runs, it is possible for aftershocks to yield large values for the energy ratio statistic, such as in the Alaskan earthquake of 21 March 1962. Also, shots may yield spurious energy peaks which probably could be eliminated by correlation weighting. At the present time, not enough shots have been run with the extended test to determine whether or not a problem of spurious peaks in the energy ratio exists.

There has been no previous discussion of pP detection using only a correlation measure among the individual seismograms. Based on previous test calculations, it is possible to state that such a test statistic would not give satisfactory results. In 37 experimental runs, using correlation alone to detect pP, misleading results occur in 13 cases. "Misleading results" mean deviations of the pP locations indicated by correlation from the presently accepted pP locations based on the best available information. Of 15 shot runs, five runs gave high enough correlation values to be mistaken for the pP phase arrival. Unexpectedly high correlation values may well be due to signal reflections from velocity discontinuities in the earth's crust. Since the velocity discontinuities, for example 8.0 km/sec to 6.34 km/sec, are not as severe as the earth-air or earth-water interface, their reflected signals will not contain as much energy, in general, as the pP or sP phases. Hence the need for an energy detection scheme.

5.4 THE EFFECT OF VELOCITY DISCONTINUITIES IN THE EARTH'S CRUST

As discussed earlier, several interfering types of signals can be adequately handled by the automatic pP test by using continental-size arrays with seismometer separations of at least 200 km. Crustal reverberations local to the event, however, are more difficult to handle. Earthquake signal reflections will occur at velocity discontinuities within the earth's crust, such as the Mohorovicic discontinuity. Evidence exists to show the presence of other discontinuities, such as the Conrad discontinuity occurring at a depth of about 10 km [9-13]. In some regions of the earth, this discontinuity may be more severe in terms of the ratio of velocities on either side than the Mohorovicic discontinuity. Reflections from these discontinuities can produce signals which are coherent, even across a continental-size array, and which, furthermore, have delay times relative to main P that are approximately the same as the P-pP delay times. Four experimental runs indicate these reflections can significantly affect the results of the pP test.

The first event affected by a reflection at a crustal-velocity discontinuity is the Kamchatka earthquake of 15 July 1963. This event was reported by the Vela Seismological Center (VSC) as a 60 km earthquake. When the shallow pP test was run, results indicated the event was a 19 km earthquake. The coda correlation discriminant showed that this event is "in the 10 to 40 km region." [Appendix A of Reference 7.] On applying the extended automatic pP test, results then confirmed the VSC depth estimate of 60 km, which was further verified by inspection of the original seismograms. This "false" peak at 19 km is probably due to the earthquake signal reflection from the Moho, which seems to be about 35 km deep in this region. The "true" peak at 60 km exhibits an energy ratio of 12 compared to an energy ratio of 2.4 at 19 km. The phases reflected from the Moho are a bit more correlated though, enough so as to fool the coda-correlation test.

The second event affected by such discontinuities is the Sea of Okhotsk earthquake of 25 August 1963. In the shallow test results, three strong peaks occurred in the η statistic, one at 12 km, a second at 22 km, and a third at 30 km. This could be accounted for by a 30 km event, with reflections from the Moho and Conrad discontinuities. A detailed picture of the crustal structure in the Okhotsk region is given by Healy [12], and the calculated delay times for such reflections are comparable with results of the automatic pP test.

The last two events affected are shots, one in the Novaya Zemlya region and one in the Kazakh region. Both events gave η peaks of greater than 1.0 when the shallow pP test was applied. The Novaya Zemlya peak occurred at a test depth of 15 km, which could have been the result of a signal path from the event down to the Conrad discontinuity (at an assumed depth of 10 km) back to the surface and on to the seismometer. The Kazakh peak occurred slightly deeper, about 18 km, which could have been caused by similar reflections with a somewhat deeper Conrad discontinuity or a somewhat slower propagation velocity. The Novaya Zemlya reflections were so correlated as to fool the coda correlation discriminant test. The Kazakh event did not do so.

Problems with these interfering reflection arrivals are inherent to the test as it is presently designed. Crustal reflection delay times should be at least roughly comparable to P-pP delay times, and since these signals come so soon after the P arrival, the relative P-pP delay times will not have a great variation across even a continental-size array. If these reflections show up consistently and strongly with shots, such as they may do in the Novaya Zemlya region, it may be necessary to study crustal structure in problem areas in order to learn what to expect if a nuclear explosion is detonated.

5.5 SUMMARY AND CONCLUSIONS

A retrospective examination of the automatic pP test has been made in order to ascertain the potential utility of the test to make reasonable depth estimates when more standard techniques based on inspection of individual seismograms may fail, to understand possible limitations on the test's performance, and to consider possible simplifications in the test's implementation. A few of the more difficult events which have been run with the automatic test have been examined in some detail.

The power of array processing designed for pP enhancement has been verified to some extent with three examples, the 6 October 1962 Ecuador earthquake, the 2 January 1964 Kamchatka earthquake, and the 21 March 1962 Alaska earthquake. For all three events, either identification of the pP (or sP) phase arrivals, or measurement of their delay times relative to P was difficult on individual seismograms, yet the pP (or sP) phase arrival was fairly easily distinguished on a sum seismogram formed to emphasize the pP (or sP) phase arrival. Use of continental-size arrays for pP processing eliminates the problem of local-to-the-seismometers crustal reverberations.

To apply array processing for pP enhancement it must be possible to add the individual signals in phase, which means relative P-pP delay times must be predictable for a given range and depth to within about 0.2 second. Although another report indicates that discrepancies of measured P-pP times from J-B predicted values by as much as 0.5 second may not be unusual, three examples have been presented which demonstrate that phasing of seismometer signals for pP enhancement offers a significant improvement in detection capability over seismometer signals summed by simply phasing on main P. Using continental-size arrays, these phasing times may be as large as 6 seconds for a 200 km event. Such large delay times can eliminate the problem of mistaking aftershocks for pP.

By reviewing previous test results, it has been found that the shallow pP test (10 to 40 km depths) performs as well

with just an energy ratio as it does with an energy ratio weighted by a correlation measure for detection of pP. However, energy ratios are comparable for 10-40 km earthquakes and shots. Correlation weighting allows the dismissal of "false" energy peaks for all but three of the twenty-two runs on which no peak should appear in the 10-40 km region. If instead of correlation weighting of the energy ratio statistic, a gross correlation measure (based on a 10 sec time window) is calculated, a simplified pP test could be used and only two classification errors in twenty-two runs would result. A similar simplification for the extended automatic pP test is not evident at this time. Thus the present design of the extended test, that is aligning for pP, and then calculating both energy ratios and correlation measures will be retained in the near future, at least until more experimental runs are performed with shots.

Velocity discontinuities in the earth's crust were shown to have affected the automatic test results on four runs, three of which were misclassified by the shallow pP test and two of which were misclassified by the coda correlation discriminant. The problems are believed to have resulted from signal reflections from either the Mohorovicic discontinuity, generally about 30 km deep, or the Conrad discontinuity, 10 km deep. Signals reflected from these discontinuities can contain significant energy and will be correlated even across continental arrays. They can also have delay times relative to main P comparable with P-pP delay times. So far this has presented an uncorrectable problem for the automatic test. One approach to resolving matters may be to study crustal structure in problem areas to learn what to expect when an event occurs. This may permit proper interpretation of test results.

APPENDIX A

WEAK NOISE ERROR IN CORRELATION TIME PICK

We assume a signal at time t' of the form

$$S(t,t') = \sqrt{E_S 2W}\ \frac{\sin 2\pi W(t-t')}{2\pi W(t-t')} \tag{A-1}$$

where W is the bandwidth of the signal.

If the value of t' is to be estimated, the situation resembles the communication problem of pulse position modulation (PPM).

Assuming the noise to be Gaussian with flat spectrum $N_0/2$ in the same band as the signal, we have from Reference [3],

$$\overline{\varepsilon_c^2} = \frac{N_0/2}{S^2} \tag{A-2}$$

where

$$S^2 = \int_{-\infty}^{\infty} \left|\frac{\partial}{\partial t'} S(t,t')\right|^2 dt \tag{A-3}$$

Thus

$$S^2 = 2WE_S \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial t'} \frac{\sin 2\pi W(t-t')}{2\pi W(t-t')}\right]^2 dt \tag{A-4}$$

This integral is better evaluated in the frequency domain, e.g., by Parseval's theorem.

The spectrum of $\psi(t)$ is

$$\Phi(f) = \begin{cases} \sqrt{E_S/2W} & |f| < W \\ 0 & \text{elsewhere} \end{cases} \tag{A-5}$$

Differentiation in time corresponds to multiplication by $j2\pi f$ in frequency, thus

$$S^2 = \frac{E_S}{2W} \int_{-W}^{W} |j2\pi f|^2 df = E_S \frac{(2\pi W)^2}{3} \tag{A-6}$$

and therefore,

$$\overline{\varepsilon_c^2} = \frac{12}{\pi^2}\left(\frac{1}{4W}\right)^2 \frac{N_0}{2E_S} \tag{A-7}$$

REFERENCES

[1] General Atronics Corporation, "First Quarterly Technical Report--Large Aperture Seismic Arrays," AD662632, September 1967.

[2] General Atronics Corporation, "Second Quarterly Technical Report -- Large Aperture Seismic Arrays," December 1967.

[3] J.M. Wozencraft and I.M. Jacobs, *Principles of Communication Engineering*, John Wiley and Sons, Inc., New York, 1965.

[4] R.M. Fano, *The Transmission of Information*, The MIT Press and John Wiley and Sons, Inc., New York, 1961.

[5] V.C. Anderson, "Digital Array Phasing," J Acous Soc Am, *32*, pp. 867-870 (1960).

[6] W.R. Remley, "Some Effects of Clipping in Array Processing," J Acous Soc Am, *39*, pp. 702-707 (1966).

[7] General Atronics Corporation, "Large Aperture Seismic Arrays," AD 657349, June 1967.

[8] Seismic Data Laboratory Quarterly Technical Summary Report (January-March 1967), AD 818674, August 1967, pp. 5-6.

[9] C.F. Richter, *Elementary Seismology*, W.H. Freeman and Company, San Francisco, 1958, Chapter 18.

[10] K.E. Bullen, *An Introduction to the Theory of Seismology*, The University Press, Cambridge, 1953, Chapter XII.

[11] B. Gutenberg, *Internal Constitution of the Earth*, Dover Publications, Inc., New York, 1951, p. 248.

[12] J.H. Healy, U.S. Geological Survey, Technical Letter Number 25, "Crust of the Earth," November 1964.

[13] J.F. Gibbs and J.C. Roller, U.S. Geological Survey, Technical Letter Number 26, "Seismic-Refraction Measurements of Crustal Structure Between Nevada Test Site and Ludlow, California," December 1964.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Generol Atronics Corporotion 1200 Eost Mermoid Lone Philodelphio, Pennsylvonio 19118	Unclossified 2b. GROUP N/A

3. REPORT TITLE

THIRD QUARTERLY TECHNICAL REPORT-
LARGE APERTURE SEISMIC ARRAYS (LASA)

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

None

5. AUTHOR(S) (First name, middle initial, last name)

None

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
Morch 1968	83	13

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF19(628)-67-C-0370 b. PROJECT NO.	ESD-TR-68-148
c. d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document hos been opproved for public releose ond sole; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Directorote of Plonning ond Technology Electronic Systems Division, AFSC, USAF L G Honscom Field, Bedford, Moss 01730

13. ABSTRACT

Four topics are discussed in this report. The first topic is the detection of underground nuclear explosions in the presence of large natural events. Previous studies of this problem have been extended and partially confirmed by applying calculations to additional data and by considering the effect of strengthening certain assumptions. The use of a LASA for estimating epicenters by beamforming and other methods is the second topic considered in this report. Recent studies of this topic have employed computer-aided simulations with controlled amounts of additive noise and travel-time anomalies to examine the relative performance of different algorithms for epicenter estimation. The third topic considered in this report is an analysis of detection and false alarm probabilities for a DIMUS array. Taking advantage of a strong analogy between this problem and a standard problem in coding theory, bounds and asymptotic expressions for these probabilities are described, and the signal-to-noise ratio loss of the DIMUS array over the corresponding analog array is analyzed as a function of several parameters. Finally, the previously developed automatic pP test has been re-examined with a view toward evaluating its performance and simplifying its implementation.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Seismology Arrays Signal Processing Seismic Discrimination Epicenter Location Earthquake Depth Estimation DIMUS						